Identification of KasA as the cellular target of an anti-tubercular scaffold

Abrahams, Katherine A; Chung, Chun-Wa; Ghidelli-Disse, Sonja; Rullas, Joaquín; Rebollo-López, María José; Gurcha, Sudagar S; Cox, Jonathan A G; Mendoza, Alfonso; Jiménez-Navarro, Elena; Martínez-Martínez, María Santos; Neu, Margarete; Shillings, Anthony; Homes, Paul; Argyrou, Argyrides; Casanueva, Ruth; Loman, Nicholas J; Moynihan, Patrick J; Lelièvre, Joël; Selenski, Carolyn; Axtman, Matthew

DOI:
10.1038/ncomms12581
10.1038/ncomms12581

Document Version
Peer reviewed version

Citation for published version (Harvard):
Abrahams, KA, Chung, C-W, Ghidelli-Disse, S, Rullas, J, Rebollo-López, MJ, Gurcha, SS, Cox, JAG, Mendoza, A, Jiménez-Navarro, E, Martínez-Martínez, MS, Neu, M, Shillings, A, Homes, P, Argyrou, A, Casanueva, R, Loman, NJ, Moynihan, PJ, Lelièvre, J, Selenski, C, Axtman, M, Kremer, L, Bantscheff, M, Angulo-Barturen, I, Izquierdo, MC, Cammack, NC, Drewes, G, Ballell, L, Barros, D, Besra, GS & Bates, RH 2016, 'Identification of KasA as the cellular target of an anti-tubercular scaffold', *Nature Communications*, vol. 7, 12581. https://doi.org/10.1038/ncomms12581, https://doi.org/10.1038/ncomms12581

Identification of KasA as the cellular target of an anti-tubercular scaffold

Katherine A. Abrahams[1], Chun-wa Chung[2], Sonja Ghidelli-Disse[3], Joaquín Rullas[4], María José Rebollo-López[4], Sudagar S. Gurcha[1], Jonathan A.G. Cox[1], Alfonso Mendoza[4], Elena Jiménez-Navarro[4], María Santos Martínez-Martínez[4], Margarete Neu[2], Anthony Shillings[2], Paul Homes[2], Argyrides Argyrou[2], Ruth Casanueva[4], Nicholas J. Loman[1], Patrick J. Moynihan[1], Joel Lelievre[4], Carolyn Selenski[5], Matthew Axtman[5], Laurent Kremer[6,7], Marcus Bantscheff[3], Iñigo Angulo-Barturen[4], Mónica Cacho Izquierdo[4], Nicholas Cammack[4], Gerard Drewes[3], Lluis Ballell[4], David Barros[4], Gurdyal S. Besra[1*], and Robert H. Bates[4*]

[1]Institute of Microbiology and Infection, School of Biosciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK; [2]GlaxoSmithKline, Gunnels Wood Road, Stevenage, SG1 2NY, UK; [3]Cellzome - a GSK Company, Meyerhofstrasse 1, 69117 Heidelberg, Germany; [4]Diseases of the Developing World, GlaxoSmithKline, Severo Ochoa 2, 28760 Tres Cantos, Madrid, Spain; [5]GlaxoSmithKline, 709 Swedeland Road, PO Box 1539, King of Prussia, PA 19406-0939; [6]Centre National de la Recherche Scientifique FRE 3689, Centre d'études d'agents Pathogènes et Biotechnologies pour la Santé, Université de Montpellier, 1919 route de Mende, 34293, Montpellier, France; [7]INSERM, CPBS, 34293 Montpellier, France.

[*]E-mail for correspondence: g.besra@bham.ac.uk (TEL: +44 121 415 8125; FAX +44 121 414 5925), robert.h.bates@gsk.com (TEL: +34 650395529).

Key words: Tuberculosis, chemical proteomics, KasA, drug discovery

Abstract

Phenotypic screens for bactericidal compounds are starting to yield promising hits against tuberculosis. In this regard, whole genome sequencing of spontaneous resistant mutants generated against an indazole sulfonamide (GSK3011724A) identifies several specific single nucleotide polymorphisms in the essential *Mycobacterium tuberculosis* β-ketoacyl synthase (*kas*) A gene. Here, this genomic-based target assignment is confirmed by biochemical assays, chemical proteomics and structural resolution of a KasA-GSK3011724A complex by X-ray crystallography. Finally, *M. tuberculosis* GSK3011724A-resistant mutants increase the *in vitro* minimum inhibitory concentration and the *in vivo* 99% effective-dose in mice, establishing *in vitro* and *in vivo* target engagement. Surprisingly, the lack of target engagement of the related β-ketoacyl synthases (FabH and KasB) suggests a different mode of inhibition when compared with other Kas inhibitors of fatty acid biosynthesis in bacteria. These results clearly identify KasA as the biological target of GSK3011724A and validate this enzyme for further drug discovery efforts against tuberculosis.

Introduction

After many years of relatively little attention, *Mycobacterium tuberculosis,* the causative agent of tuberculosis (TB), has re-emerged as a priority in the area of neglected diseases[1]. The standard six-month treatment for TB has remained essentially unchanged over several decades[2]. Unfortunately, poor patient compliance and other factors have led to an increasing prevalence of drug resistance. In 2013, the World Health Organization (WHO) recorded approximately 480,000 new cases of multi-drug resistant (MDR)-TB, resistant to both the front-line drugs, isoniazid (INH) and rifampicin (RIF)[3]. Without these two front-line drugs, MDR-TB generally requires twenty-four months of treatment with a variety of second-line antibiotics, which are poorly tolerated. Due to this continued spread of MDR and extensively-drug resistant (XDR) forms of *M. tuberculosis,* the WHO has declared TB as a world health emergency[3]. Therefore, new drugs that can complement existing front-line treatment regimens are urgently required.

In 2013 and updated in 2015, GlaxoSmithKline published a set of 228 small molecule hits from two phenotypic screening campaigns against *M. tuberculosis*[4,5]. Of those 228 compounds, a number have been explored *via* medicinal chemistry for potential optimisation to drug leads and clinical candidates. Phenotypic screening also offers the opportunity to identify novel biological targets by molecular mode of action (MoA) studies, which is generally achieved through a variety of omics-based technologies[6-10]. GSK3011724A is a small molecule inhibitor from the aforementioned 228 phenotypic screening hits[4,5]. In this study, we provide an initial profiling of GSK3011724A as a suitable starting point for 'hit' to 'lead' drug

development based on its *in vitro* and *in vivo* characteristics. Additionally, through MoA studies, we identify the *M. tuberculosis* cellular target of GSK3011724A to be KasA, a β-ketoacyl synthase. Target assignment reveals specificity of GSK3011724A for KasA with a binding site distinct from other known Kas inhibitors. This different MoA creates new potential for this recognised target in future TB drug discovery efforts.

Results

In vitro profiling of GSK3011724A

From the outset, GSK3011724A, an indazole sulfonamide (**Fig. 1a**), represented an attractive compound for early stage drug discovery based on its anti-mycobacterial potency, small size and moderate lipophilicity (**Table 1**). Further profiling *in vitro* using standard assays, such as the hERG ion channel and cytochrome P450 isoforms supported this view (**Table 1**). Sensitivity testing of 18 Gram-positive and Gram-negative bacterial species with GSK3011724A demonstrated selectivity of this compound for *M. tuberculosis*. Additionally, GSK3011724A showed negligible activity against a panel of unrelated proteins (**Supplementary Table 1**). These desirable characteristics demonstrated the clear potential of GSK3011724A as a starting point for medicinal chemistry optimisation. A description of this optimisation effort along with a thorough discussion of the drug development potential and liabilities of this series will be the subject of a future publication.

In vivo profiling of GSK3011724A

GSK3011724A was progressed into pharmacokinetic (PK) and efficacy experiments to generate an *in vivo* profile. Although doses above 400 $mg.kg^{-1}$ were not well tolerated by mice, multiple days of dosing up to 300 $mg.kg^{-1}$ once daily proceeded with no weight loss or other adverse effects. The *in vivo* PK of GSK3011724A in mice was disproportional based on dosing. At low doses, clearance close to the liver blood flow rate was observed (in agreement with the *in vitro* clearance (Cli) in mouse microsomes), but as the doses were increased, the maximum concentration (Cmax) and area under the curve (AUC) values grew disproportionally (**Table 2**). This observation suggests a saturation of the mechanism of clearance, and may help to explain the reduction in tolerability between 300 and 400 $mg.kg^{-1}$ dosing.

The first stage of efficacy testing involved an acute model of infection[11]. This murine model involved infecting mice with a high inoculum (10^5 colony forming units (CFU)) of *M. tuberculosis* H37Rv wild-type and beginning treatment the next day. GSK3011724A (and INH as a positive control) were administered once daily for 8 days, while the bacilli were in the exponential growth phase. Despite exhibiting a bacteriostatic effect *in vitro* (**Supplementary Fig. 1**), GSK3011724A demonstrated a significant cidal effect in this murine model with an approximate 3.5 log CFU reduction relative to untreated controls at 200 $mg.kg^{-1}$ (**Fig. 1b**). The effects of GSK3011724A are similar to linezolid, which has been shown to be bacteriostatic *in vitro*[12], but exhibits cidality *in vivo*[11]. Importantly, a clear dose response was observed, with the ED_{99} (dose required to reduce bacterial load by 99 %) determined as 38 $mg.kg^{-1}$. This level of ED_{99} is only moderately higher than most of the gold standard TB drugs[11].

GSK3011724A (and INH as a positive control) was also evaluated in a chronic infection assay in which mice were infected with a lower inoculum of *M. tuberculosis* H37Rv (10^2 CFU) and left untreated for 6 weeks, allowing the bacilli to reach a steady state. GSK3011724A was then dosed daily for two months. The dose response curve in this chronic assay was shifted to higher doses than in the acute model, but the targeted >2 log CFU reduction was achieved at 100 $mg.kg^{-1}$ for GSK3011724A, producing a 2.4 log CFU reduction (**Fig. 1c**). These data clearly establish GSK3011724A as an active compound in both the acute and chronic *in vivo* murine models of *M. tuberculosis* infection. Importantly, if GSK3011724A were demonstrated to inhibit an unexploited potential antimicrobial target, these data would provide a significant level of validation and confidence for further exploration of that target for TB drug discovery.

GSK3011724A target identification

A fundamental strategy in drug discovery is establishing the MoA of inhibitory compounds. Following the identification and validation of the molecular target, target-specific optimisation of the compound can be pursued in order to improve efficacy and reduce toxicity. Due to the recent successes of utilising whole genome sequencing (WGS) of spontaneous resistant mutants as a primary step in the elucidation of the target of phenotypic hits[6,7,9], this methodology was used to establish the target of GSK3011724A.

The minimum inhibitory concentration (MIC) of GSK3011724A in *Mycobacterium bovis* BCG was determined to be 0.5 μM. Spontaneous resistant mutants were initially generated using *M. bovis* BCG at 5 x, 10 x and 20 x the MIC of GSK3011724A, with frequencies of resistance (FoR) of 12 x 10^{-8}, 3 x 10^{-8} and 2 x 10^{-8}, respectively. The FoR was subsequently determined against *M. tuberculosis* at 2.5 μM (10 x MIC of GSK3011724A on solid media, 0.25 μM) giving a frequency of 9.5 x 10^{-7}. This FoR is slightly higher than normally desired but is lower than that of INH.

From the total of seventeen *M. bovis* BCG spontaneous resistant mutants generated, eight were selected for WGS. Six of the eight *M. bovis* BCG mutants were found to possess a number of high-quality (statistically relevant) single nucleotide polymorphisms (SNP) compared to the sequenced wild-type reference strain (Genbank accession number NC_008769.1), all of which located to the gene annotated *kasA*, encoding an essential β-ketoacyl synthase involved in mycolic acid biosynthesis[13] (**Table 3**). The data offered the first evidence that KasA is the target of GSK3011724A, providing a starting point for validation studies. Subsequently, instead of using WGS, the *kasA* gene from twelve isolated *M. tuberculosis* resistant mutants was specifically sequenced following amplification by PCR. The data confirmed the results from *M. bovis* BCG, with eleven of the twelve mutants showing SNPs in *kasA* (**Table 3**). The WGS results of the two remaining *M. bovis* BCG mutants did not reveal SNPs of a high frequency; no gene contained a SNP of greater than 50 % frequency, corresponding to the percentage of SNP in the cell population from which the genomic DNA was prepared.

KasA target validation

Following the target identification of KasA by WGS, a variety of techniques were utilised to confirm its role as the biological target of GSK3011724A. Firstly, the inhibition of mycolic acid biosynthesis by GSK3011724A was investigated. *M. bovis* BCG was labelled using [^{14}C]-acetate and treated with increasing concentrations of GSK3011724A, with INH and the KasA inhibitor thiolactomycin (TLM) used as positive controls. Following drug treatment, the total fatty acid methyl esters (FAMEs) and α-mycolate and keto-mycolic acid methyl esters (α- and k-MAMEs) were extracted and analysed by autoradiography thin-layer chromatography (TLC). As shown in **Figure 2a** (left panel) there was a clear visual dose-dependent reduction in the production of α-MAMEs and k-MAMEs upon treatment with GSK3011724A. This characteristic profile of total MAMEs inhibition and FAMEs accumulation mirrors the known fatty acid synthase (FAS)-II inhibitors TLM and INH (**Fig. 2a**, middle panel). When spots were excised from the TLC and quantified by scintillation counting (**Table 4**), GSK3011724A at 4 x MIC in comparison to the non-treated control produced a 55.2 % and 42.3 % reduction in α-MAMEs and k-MAMEs, respectively. In contrast, quantification of the levels of FAMEs increased by 186.5 % at 4 x MIC. When compared to the known KasA inhibitor TLM, GSK3011724A at 1 x MIC afforded a similar level of inhibition of α-MAMEs and k-MAMEs, and an increase in FAMEs synthesis, respectively (**Table 4**). A slightly less pronounced effect at their relative MICs was observed for INH in comparison to GSK3011724A (**Table 4)**. The inhibition of mycolic acid biosynthesis by GSK3011724A was even more apparent when cell wall bound α- and k-MAMEs were analysed, with an almost visual complete inhibition in a dose-dependant manner, which was similar to INH and TLM (**Fig. 2a**, right panel). To investigate whether GSK3011724A affects the

synthesis of other mycobacterial lipids, addressing the potential of additional cell wall associated targets, the total cellular lipids (and fractionated apolar lipids and polar phospholipids) were extracted and analysed following drug treatment. Firstly, a decrease in trehalose dimycolate (TDM), trehalose monomycolate (TMM) and glycerol monomycolate (GroM) was observed with GSK3011724A (and TLM and INH), consistent with inhibition of mycolic acid biosynthesis (**Fig. 2b**, **Table 4**, **and Supplementary Fig. 2a**). Secondly, no significant differences were observed in lipid composition when apolar lipids and polar phospholipids were analysed by autoradiography-TLC (**Supplementary Fig. 2a, b**), suggesting that GSK3011724A does not target other lipid biosynthetic pathways, and is specific for KasA.

To further corroborate that GSK3011724A inhibits mycolic acid biosynthesis and, more specifically, targets KasA, the impact on the MIC was investigated using strains of *M. bovis* BCG over-expressing components of FAS-II. The mycobacterial expression vectors, pMV261 containing *M. tuberculosis kasA*[14], *kasB*[14], *fabH*[15], *inhA*[16], *hadABC*[17] and *mabA*[16], were electroporated into *M. bovis* BCG and growth analysed at a concentration range with respect to the MIC of GSK3011724A (**Fig. 2c**). *M. bovis* BCG containing the pMV261-empty vector exhibited no change in sensitivity to GSK3011724A compared to the wild-type untransformed strain, with an MIC of 0.5 μM (**Fig. 2c**). This was also apparent for the pMV261 constructs containing *kasB*, *inhA*, *fabH*, *hadABC* and *mabA* (**Fig. 2c**). However, ample growth of the KasA over-expressor strain was observed at 4 μM, indicating an increase in resistance, and a MIC shift of >8 x (**Fig. 2c**), providing further evidence to support KasA as the cellular target of GSK3011724A.

Previous studies have reported KasA activity in an *in vitro* condensing assay using purified recombinant proteins[18]. This assay was replicated to demonstrate the specific inhibition of KasA by GSK3011724A. Initially, FabD was used to generate [^{14}C]-malonyl-AcpM from holo-AcpM and [^{14}C]-malonyl-CoA. KasA, in the absence or presence of GSK3011724A, elongated C_{16}-AcpM in a condensation reaction with [^{14}C]-malonyl-AcpM. Treatment with 1 μM of GSK3011724A provided a 58.5% inhibition of KasA activity (based on triplicate data, **Supplementary Fig. 3**).

To investigate whether GSK3011724A exerts its effect by directly binding to KasA, we employed a chemoproteomics strategy[19,20]. A tagged analogue (**1**) with an MIC of 49 μM was generated, which was covalently linked to Sepharose beads (**Fig. 3a**). The tagged analogue-(**1**)-derivatised beads were incubated with *M. bovis* BCG extracts, in the absence or presence of an excess of unbound GSK3011724A. In this competition-binding format, target proteins would be expected to bind to the beads predominantly in the absence of excess competing compound. After washing of the beads, bound proteins were digested with trypsin and subjected to quantitative mass spectrometry. Among more than 2,000 proteins identified, only three proteins were efficiently competed off the beads by excess GSK3011724A: KasA, and, to a lesser degree, the non-essential polyketide synthases Pks10 and Pks11[21,22] (**Fig. 3a, Supplementary Data 1** and **2**). The experiment was repeated in a concentration-dependent format to determine half-maximal inhibition (IC_{50}) values that refer to the concentration of compound (GSK3011724A) required to competitively block 50 % of the target (i.e. KasA) from binding to the beads (**Fig. 3b, Supplementary Data 3** and **4**). These IC_{50} values are a measure of target affinity, but are also affected by the affinity of the target for the bead-immobilised ligand. The latter effect can be deduced by measuring

the depletion of the target proteins (KasA, Pks10 and Pks11) by the beads. Thus, apparent dissociation constants (K_d^{app}) can be determined, which are largely independent from the bead ligand[19,23]. The K_d^{app} value was determined as 9 nM for KasA, suggesting a high level of affinity for the enzyme. In excellent agreement with these chemoproteomic affinities, direct binding of GSK3011724A to purified KasA was also determined using Surface Plasmon Resonance (SPR) and found to be in the 10-20 nM range (**Supplementary Fig. 4**). A much weaker apparent dissociation constant was observed for both Pks10 and Pks11 (K_d^{app} 1.4 μM). In addition, we compared the results of the active compound GSK3011724A with a structurally related inactive analogue **2** (MIC > 125 μM). The K_d^{app} for this inactive molecule was determined to be 4.5 μM for KasA, which corresponds to a ~490 fold window for KasA between active and inactive compound. There was no binding observed for Pks10 and Pks11 by the inactive analogue **2** (**Fig. 3b, Supplementary Data 5**).

In order to evaluate the potential of GSK3011724A for interaction with human proteins, which could represent off-targets relevant for drug safety, we performed a similar set of chemoproteomics experiments with a human protein extract, generated by combining detergent (NP40) lysates from HEK293 cells, K562 cells and placenta tissue (**Supplementary Data 6**). Only a single protein showed moderate competition by GSK3011724A, NQO2, with an apparent dissociation constant of 4-6 μM. NQO2 is an off-target of many small molecules drugs[24]. Other potential off-targets may exist that are not expressed in the cells and tissues used for the profiling, and the compound may exhibit binding modes, which are inconsistent with its immobilisation. However, taken together, the chemoproteomics data demonstrate that GSK3011724A is a highly

potent and selective inhibitor of KasA with no discernible off-targets in the above human cell lines and placenta tissue.

Analysis of the KasA-GSK3011724A co-crystal structure

Curious to understand the molecular details of GSK3011724A binding to KasA, a 2.13 Å co-crystal structure of the dimeric KasA-GSK3011724A complex was solved (**Table 5**). All literature inhibitors to date, such as TLM, reside in the malonyl substrate pocket close to the catalytic Cys171 residue[25,26] (**Fig. 4a**). However, computational docking studies to place GSK3011724A into this site produced puzzling binding modes, which did not explain the roles of the key molecular features of GSK3011724A. The co-crystal structure reveals that, in contrast to TLM, GSK3011724A uniquely binds within the large acyl channel that normally accommodates the growing meromycolic acid chain, unexpectedly trapping the open conformation of the enzyme (**Fig. 4a**). GSK3011724A sits at the branch point of the acyl channel, a feature thought to confer specificity for "kinked" unsaturated substrates. Within this site, density for a long linear molecule, modelled as polyethylene glycol (PEG), is also consistently present and perfectly complements the space between the compound and the protein (**Fig. 4b**). This sandwich of PEG and GSK3011724A enables the large acyl cavity to be filled, occupying the volume of the branched phospholipid chains previously observed within this pocket[25].

From the co-crystal structure, the structure-activity relationship of GSK3011724A is evident (**Fig. 4b, c**). The indazole ring lies across the flat hydrophobic surface formed by residues such as Gly200 and Pro201, and the amide of the sulfonamide makes a

key hydrogen bond with Glu199. This anchors the ligand in an ideal position, allowing the kink (introduced by the sulfonamide linkage) to place the butyl chain into an orthogonal narrow lipophilic channel lined by residues such as Ile347, Ile202 and Phe239.

The resistance-conferring mutations observed in **Table 3** can be readily rationalised by the inhibitor binding site and MoA of GSK3011724A (**Fig. 5a**). Pro201 and Pro206 both lie within the acyl site and are in direct contact with the ligand (**Fig. 5b**). Exchange of proline for a hydrophilic serine residue (Pro201Ser) disfavours interactions with the aromatic indazole ring, whereas the Pro206Leu substitution introduces steric crowding of the ligand. The remaining mutations are found away from the inhibitor site, in regions likely to influence the conformational equilibrium and ease of transition between the closed and open state required for GSK3011724A inhibition. The *α5* to *α6* helix-turn-helix (HTH) arm of one KasA monomer, comprising of residues 115-145, must slide past its dimer counterpart (*α′5*-*α′6*) and the *α′2* helix as the acyl cavity opens. Mutations Leu128Pro, Val137Ala, Met146Val and Leu205Pro are either within the HTH motif or the *α2* helix and undergo substantial movement to the open form (**Fig. 5c**). Met277 also lies at the dimer interface, however, its position is fixed between the two states. This residue could be considered as the 'ball' in a 'ball and socket' joint around which the *α5*-*α6* HTH pivots (**Fig. 5d**). Reducing the size of Met277 to valine may alter the movement and position of the HTH. Similarly, Thr114 sits at another fixed pivot point, the start of the first helix of the HTH (**Fig. 5d**). The positioning of Thr114 breaks the β-strand that proceeds it and anchors the HTH arm to its opposite β-strand partner Val198 *via* two hydrogen bonds (through a carbonyl-backbone interaction and by the interaction

of the threonine side chain hydroxyl to the carbonyl of Val198) (**Fig. 5e**). The Thr114Ser resistance conferring mutation would introduce more flexibility and may reduce the anchoring stability needed in this position.

In summary, these crystallographic insights complete a consistent picture of the molecular MoA of this inhibitor and provide a platform for rational optimisation of the GSK3011724A scaffold as well as *de novo* structure-based drug design. Excitingly, this offers a fresh opportunity to target KasA at a site distinct from previous inhibitors and one that is able to achieve selectivity over other related β-ketoacyl synthases (FabH and KasB) involved in fatty acid biosynthesis, as key residues required for affinity are not conserved (**Fig. 5f**, **g**).

In vivo target engagement of GSK3011724A and KasA

Although the evidence presented above provided a high level of confidence that KasA was the true biological target of GSK3011724A, there was still a need to confirm that target engagement was responsible for the potent *in vivo* activity. A selection of the *M. tuberculosis* spontaneous resistant mutants isolated against GSK3011724A were tested for their growth profile in mice using the acute model[11] (**Fig. 1d**). While the two mutants showed an attenuated growth rate in untreated animals, their response to INH treatment remained essentially unchanged. Given the close relationship between KasA and InhA, the INH data were particularly significant, confirming that the observed resistance *in vivo* was not merely an artefact of the limited growth rate. In contrast with INH, the response with GSK3011724A against both strains with mutations mapping to *kasA* showed clear signs of resistance (**Fig. 1d**). Mutation

Met277Thr appeared to impart complete resistance up to the maximum dose tested (140 $mg.kg^{-1}$), while mutation Pro201Ser gave a lesser response at 100 $mg.kg^{-1}$. These results support the MIC shifts observed *in vitro* and furthermore provide a critical link between the target identification and *in vivo* validation of KasA.

Discussion

On-going efforts to combat drug-resistant TB have taken many forms including the re-purposing of broad spectrum antibacterials, target-based programs on mycobacterial enzymes, efforts to optimise or re-invent known TB drugs (like INH and RIF), and phenotypic screening approaches, which have all been widely reported[4,27-32]. In this work, through a variety of *in vitro* and *in vivo* experiments, a new chemical scaffold, exemplified by GSK3011724A, has been identified to specifically target an integral component of mycolic acid biosynthesis, KasA[18]. Mycolic acids are unique and fundamental components of the mycobacterial cell wall and KasA is essential in *M. tuberculosis*[13].

Mycolic acid biosynthesis involves two distinct fatty acid synthesis pathways. The FAS-I system is required for *de novo* fatty acid synthesis, where a single, multifunctional polypeptide generates short chain fatty acyl-CoA esters. FabH, a β-ketoacyl ACP synthase, forms a pivotal link between FAS-I and FAS-II, condensing C_{14}-CoA (generated by FAS-I) and malonyl-AcpM producing C_{16}-AcpM[15]. This product is channelled to KasA of the FAS-II system[33]. The FAS-II system is comprised of four enzymes acting in a consecutive cycle: KasA and KasB, condensing enzymes[18]; MabA, a keto-reductase[34]; HadABC, a dehydratase[35]; and

InhA, an enoyl-reductase[36]. The FAS-II system enables fatty acid elongation leading to meromycolic acids (C_{56}), which are then condensed with C_{26}-CoA (from FAS-I) by the polyketide synthase Pks13[37,38], followed by reduction, culminating in the production of mature mycolic acids[39]. The lipidomics experiments showing accumulation of FAMEs and depletion of MAMEs confirms that GSK3011724A specifically inhibits FAS-II and mycolic acid biosynthesis (**Fig. 2a**, **b, Table 4** and **Supplementary Fig. 2**).

Inhibitors of KasA have been reported in the literature[14,40,41]. Most notable among these is TLM, which is known to inhibit all three mycobacterial Kas enzymes: KasA, KasB, and FabH[14,15]. Interestingly, GSK3011724A, unlike TLM and other Kas-inhibitors, whether against *M. tuberculosis* or other bacteria (*via* FabH, FabB and FabF), displayed unique specificity, targeting only KasA (**Fig. 2c**). An explanation of this finding comes from close inspection of the GSK3011724A-KasA complex crystal structure, where specific changes in the acyl pockets of the related enzymes disfavour the unique binding mode of GSK3011724A **(Fig. 4** and **5)**. For example, the Gly200Arg and Pro201Thr changes from KasA to KasB no longer allow the indazole ring to sit favourably in the acyl pocket of KasB **(Fig. 5f)**. This offers an explanation as to why the WGS data did not result in the identification of SNPs in KasB and why the resistance mutations in KasA often map to points of variation amongst the Kas enzymes.

Importantly, GSK3011724A, despite representing an un-optimised screening hit, gives at least one order of magnitude greater potency relative to the most studied KasA inhibitor, TLM, and its favourable PK properties allowed for the critical *in vivo*

394 experiments described in this work. The data obtained from murine assays (both acute
395 and chronic infection) exemplify the credible potential of KasA inhibitors to give
396 significant efficacy and, to our knowledge, offer the first *in vivo* validation of KasA as
397 a drug target (**Fig. 1d**). Together with the promising drug-like profile of
398 GSK3011724A, these data provide significant confidence for the future exploration of
399 KasA as a drug target, inhibitors of which could become key players in the
400 development of new anti-tubercular drugs and will be explored in a future publication.

Materials and Methods

Synthetic and characterisation details for the compounds and the SPR experiment described herein can be found in the **Supplementary Methods**. All animal studies were ethically reviewed and carried out in accordance with European Directive 2010/63/EU and the GSK Policy on the Care, Welfare and Treatment of Animals. The human biological samples were sourced ethically and their research use was in accordance with the terms of informed consent.

Assessment of acute and chronic efficacy in murine TB models

INH was purchased from Sigma Aldrich and prepared freshly in distilled water. GSK3011724A was prepared freshly in 1 % aqueous methylcellulose. The assessment of the chronic and acute efficacy in murine TB models was performed using specific pathogen-free, 8-10 week-old female C57BL/6 mice purchased from Harlan Laboratories and allowed to acclimate for one week and kept under controlled conditions in a P3 high security facility with unlimited sterile food and water.

In the acute model[11], mice were intratracheally infected with *M. tuberculosis* H37Rv wild-type (H37Rv WT) 100,000 CFU for each mouse, and lungs harvested on day 9. GSK3011724A and INH were administered daily for 8 consecutive days, starting on day 1 after infection. In the chronic model[42], mice (n = 2 mice at each dose level) were intratracheally infected with 100 CFU for each mouse; INH or GSK3011724A was administered daily for 8 consecutive weeks, starting 6 weeks after infection. Lungs were harvested 24 h after the last administration in both assays. All lung lobes

were aseptically removed, homogenised, and frozen. Homogenates were unfrozen and plated in 10 % OADC-7H11 medium supplemented with activated charcoal (0.4 %) and grown for 18-25 days at 37°C. Non-linear fitting was performed with the dose response data (logCFU versus dose) and the dose in $mg.kg^{-1}$ that reduced lung bacterial burden by 99 % with respect to untreated mice was estimated (ED_{99}). Mice were supervised every day under a protocol paying attention to weight loss, apparent good health (bristled hair and wounded skin) and behaviour (signs of aggressiveness or isolation). Animals were euthanised by CO_2 inhalation.

MIC determination and resistant mutant generation

M. bovis BCG strain Pasteur and derivatives were cultured at 37°C and 5 % CO_2 in static liquid or solid medium. Liquid medium contained Middlebrook 7H9 (Difco) supplemented with 0.05 % (v/v) Tween-80, 10 % (v/v) Middlebrook ADC and 0.25 % (v/v) glycerol. Solid medium contained Middlebrook 7H11 agar (Difco) with 10 % (v/v) Middlebrook OADC and 0.5 % (v/v) glycerol. Where applicable, 25 $\mu g.mL^{-1}$ Kanamycin was added to the liquid or solid media to select for mycobacterial plasmids. The constructs pMV261, pMV261-Mt-*kasA*, pMV261-Mt-*kasB*, pMV261-Mt-*fabH*, pMV261-Mt-*inhA*, pMV261-Mt-*hadABC* and pMV261-Mt-*mabA* were electroporated into *M. bovis* BCG[14-17]. Wild-type *M. bovis* BCG electrocompetent cells were prepared by pelleting a mid-log culture and washing with decreasing volumes of ice-cold 10 % (v/v) glycerol. The cells were incubated on ice with 1 µg plasmid DNA before being transferred to a 0.1 cm electrode-gap electroporation cuvette and subjected to a single pulse of 1.8 kV. Cells were recovered in liquid

media overnight at 37°C and selected on solid medium containing the appropriate antibiotic.

The MIC of GSK3011724A was determined by plating 10^4, 10^3, 10^2, and 10^1 cells from a mid-log culture of *M. bovis* BCG on solid medium containing increasing concentrations of compound in a dose response format. The MIC was defined as the concentration of compound that caused complete inhibition of bacterial growth. *M. bovis* BCG and *M. tuberculosis* spontaneous resistant mutants were generated by plating 10^8 cells from a mid-log phase culture on solid media containing either 5 x, 10 x or 20 x MIC of GSK3011724A. Potentially resistant colonies were inoculated into liquid media, cultured to mid-log growth phase, and selected on solid media containing 5 x MIC of GSK3011724A to confirm phenotypic resistance. The MIC of the resistant *M. tuberculosis* mutant strains against GSK3011724A was determined either using the MABA resazurin assay[43] or by serial dilution and agar plating.

Sequencing of resistant mutants

Wild-type *M. bovis* BCG and the *M. bovis* BCG GSK3011724A resistant mutants were characterised by WGS[6,9]. Briefly, purified genomic DNA was prepared for sequencing using the Nextera DNA Sample Preparation Kit (Illumina). The DNA libraries were purified and quantified using Agencourt AMPure XP beads (Beckman Coulter Genomics) and Quant-iT PicoGreen dsDNA kit (Life Technologies), respectively. Fragment sizes were determined using an Agilent Technologies 2100 Bioanalyzer with a High Sensitivity DNA chip. Following the MiSeq preparation guide, the libraries were sequenced on a MiSeq Benchtop Sequencer using the MiSeq

475 Reagent Kit v2, 300 cycles. Reads were aligned to the reference genome *M. bovis*
476 BCG Pasteur 1173P2 (accession: NC_008769.1).
477
478 The *kasA* gene (Rv2245) from *M. tuberculosis* GSK3011724A resistant mutants was
479 sequenced specifically. The gene was amplified by PCR using 0.5 μM of the flanking
480 primers 5´- aggacaagtacggcgtcaag (forward) and 5´- gtaaccagctccgtcattg (reverse).
481 The amplification was performed with 0.02 $U.\mu L^{-1}$ of KOD Xtreme™ Hot Start DNA
482 Polymerase (Merck Millipore) and 0.4 mM each of dNTP. PCR conditions were: 2
483 min at 94°C, followed by 30 cycles of 98°C for 10 s, 66°C for 60 s and 68°C for 10 s.
484 The PCR product was purified from the band of a 0.8 % agarose gel (Seakem GTG
485 Agarose, Lonza) with Illustra™ GFX™ PCR DNA and Gel Band Purification Kit,
486 GE Healthcare. The purified product was used for *kasA* automated sequencing (using
487 the primers: Forward 1, cgaagattgagtcggagaac; Reverse 1, cttccatatcggtccgactc;
488 Forward 2, gtcaagatcggcggtcac; Reverse 2, cagacttcggcgcgtaca; Forward 3,
489 gtgatcggtctgcagcttg; Reverse 3, ctcctccgtctcgatgag; Forward 4, ctcatcgagacggaggag;
490 Reverse 4, cgttgggcatgatcatctg; Forward 5, tgtacgcgccgaagtctg; Reverse 5,
491 gatcccacttggtgacgaac). DNA sequencing reactions were performed with a BigDye
492 terminator V3.1 cycle sequencing kit (Applied Biosystems, Inc.).
493
494 **Synthesis of FAMEs and MAMEs**
495
496 The whole cell effect of GSK3011724A was studied by treating *M. bovis* BCG
497 cultures (10 mL) at an OD_{600nm} of 0.4-0.6 with a dose dependent increase in drug for
498 20 h before labelling using 1 $\mu Ci.mL^{-1}$ [1-^{14}C]sodium acetate (37 MBq, PerkinElmer)
499 for a further 24 h at 37°C . The total FAMEs and MAMEs were extracted[14,44]. Briefly,

cells were pelleted and incubated overnight in 2 mL of 5 % tetrabutylammonium hydroxide at 100°C. The following day, 4 mL of dichloromethane was added with 300 µL of iodomethane and 2 mL of water and mixed for 30 min. The reaction was centrifuged and the upper aqueous layer discarded. Water (3 mL) was added to the lower organic layer, mixed and centrifuged as before and repeated once more. The organic layer was evaporated to dryness and the methyl esters re-dissolved in diethyl-ether (4 mL) and transferred to a fresh tube. The diethyl-ether was evaporated and 200 µL of dichloromethane used to re-dissolve the extracted methyl esters. The total FAMEs and MAMEs were analysed by TLC, using equal counts (cpm) and exposed to Kodak X-Omat film. Quantification of labelled FAMEs and MAMEs was determined by excising spots directly from the TLC plates and subjecting them to scintillation counting using 10 mL of EcoScintA.

Extraction of cell wall bound MAMEs and lipids

M. bovis BCG was drug-treated and labelled as described above. The cell wall bound MAMEs, apolar and polar phospholipids were extracted and analysed[44]. Briefly, the cell pellet was extracted 4 times using 4 mL of chloroform:methanol:water (10:10:3, v/v/v), retaining the delipidated cell pellet (for analysis of cell wall bound mycolic acids) and collecting the solvent extract sequentially, which was combined and dried. To the dried extract, 4 mL of chloroform:methanol:water (10:10:3, v/v/v) was added, followed by chloroform (1.75 mL) and water (0.75 mL), and the entire mixture centrifuged and the lower organic layer recovered. The lower organic layer was washed twice using chloroform:methanol:water (2 mL, 3:47:48, v/v/v) and dried to provide a total lipid extract, which was re-dissolved in chloroform:methanol (2:1, v/v)

525 and an aliquot subjected to scintillation counting using 10 mL of EcoScintA. The total
526 lipid extract was further partitioned between the phases arising from methanol:0.3 %
527 NaCl (2 mL, 100:10, v/v) and 2 mL of petroleum-ether (60-80^{o}C). The entire contents
528 were mixed on a blood rotor, centrifuged and the upper layer collected. The lower
529 layer was re-extracted using 2 mL of petroleum ether (60-80^{o}C). The combined
530 petroleum-ether layers were evaporated to afford the crude apolar lipids. To the lower
531 organic layer, 2.3 mL of chloroform:methanol:0.3 % NaCl (50:100:40, v/v/v) was
532 added, followed by 750 µL of chloroform:methanol:0.3 % NaCl (50:100:40) and a
533 further 1.3 mL of chloroform and 1.3 mL 0.3 % NaCl. The entire contents were
534 mixed, centrifuged and the lower layer recovered and dried to afford the crude polar
535 lipids. The apolar and polar lipids were re-dissolved in chloroform:methanol (2:1, v/v)
536 and an aliquot subjected to scintillation counting using 10 mL of EcoScintA. The
537 total lipid extracts and apolar/polar lipids were analysed by TLC in the following
538 solvent systems using equal counts (as stated) before being exposed to Kodak X-
539 Omat film. Apolar lipids were resolved using three solvent systems A: first direction,
540 petroleum-ether (60:80^{o}C)):ethyl acetate (98:2, v/v, thrice); second direction,
541 petroleum-ether (60:80^{o}C):acetone (98:2, v/v). Solvent system B: first direction,
542 petroleum-ether (60:80^{o}C):acetone (92:8, v/v, thrice); second direction,
543 toluene:acetone (95:5, v/v). Solvent system C: first direction, chloroform:methanol
544 (96:4, v/v); second direction, toluene: acetone (80:20, v/v). The polar lipids were
545 analysed using chloroform:methanol:ammonium hydroxide:water (65:25:0.5:3.6,
546 v/v/v). The total lipid extract was analysed using chloroform:methanol:concentrated
547 ammonium hydroxide (80:20:2, v/v/v) to reveal TDM and TMM. The recovered
548 delipidated cells were used to analyse cell wall bound mycolic acids following release
549 using 5 % tetrabutylammonium hydroxide at 100°C and methylation as described

above for total FAMEs and MAMEs. The recovered cell wall bound MAMEs were analysed by TLC, using an equal aliquot (5 %) and exposed to Kodak X-Omat film.

Original scans of all TLCs are shown in **Supplementary Figure 5**.

KasA activity assay

Recombinant FabD, holo-AcpM and C_{16}-AcpM were over-expressed in C41 (DE3) *E. coli* cells from pET28a-*fabD* and pET28a-*acpM* and purified[45]. Briefly, cells were resuspended in buffer (50 mM potassium phosphate, pH 7.5, 0.5 M NaCl and 10 mM (AcpM) / 25 mM (FabD) imidazole) containing DNAse, Complete protease inhibitor-cocktail tablets (Roche) and 0.1 $mg.mL^{-1}$ lysozyme. Cells were disrupted by 6 passes through a French Press and the clarified lysate was loaded onto a pre-equilibrated (with buffer) Ni^{2+}-charged 1 mL His-Trap column. A step gradient of imidazole (50-1000 mM) was used to wash and elute the recombinant protein. Recombinant AcpM was dialysed into 0.1 M Tris pH 7.5, 500 mM NaCl and loaded onto a column containing 1 mL Thiopropyl-Sepharose 6B. Acyl-AcpM was collected in the flow through. Holo-AcpM was eluted with increasing concentration of β-mercaptoethanol (5-100 mM). All purified enzymes were dialysed firstly against 50 mM Tris pH 7.5, 50 mM NaCl, 10 % (v/v) glycerol, 2 mM EDTA and secondly against the same buffer without EDTA.

The *E. coli* expression plasmid, pET28a-*kasA*[18], was transformed into *E. coli* BL21 (DE3). A resulting single colony was used to inoculate an overnight culture, which was subsequently used to inoculate 1 L LB broth, 1 % (w/v) glucose, 50 $\mu g.mL^{-1}$

575 Kanamycin. The culture was incubated at 37°C, 180 rpm, until OD_{600nm} reached 0.4-
576 0.6. The culture was cooled to 16°C and induced with 1 mM IPTG. Growth was
577 continued for 20 h at 16°C, 180 rpm, and then the cells were harvested by
578 centrifugation. The cell pellet was resuspended in buffer (50 mM potassium
579 phosphate, pH 7.45, 0.5 M NaCl and 10 mM imidazole) containing DNAse, RNAse,
580 Complete protease inhibitor-cocktail tablets (Roche) and 0.1 $mg.mL^{-1}$ lysozyme.
581 Cells were disrupted by 6 passes through a French Press at 1 ksi and centrifuged at
582 15,000 rpm, 40 min at 4°C. The supernatant was loaded onto a pre-equilibrated (with
583 buffer) Ni^{2+}-charged 1 mL His-Trap column. The column was washed extensively
584 using buffer and the proteins were eluted with a step-wise gradient of imidazole (50,
585 100, 150, 200, 350 and 500 mM). SDS-PAGE was used to detect the presence of
586 purified KasA, which was dialysed against 50 mM Tris.HCl, pH 7.5, 300 mM NaCl,
587 10 % (v/v) glycerol and stored at -20°C.
588
589 The assay for mycobacterial KasA activity was performed[18]. Assay components were
590 mixed together in a batch fashion, before equally dividing according to the number of
591 assays performed. The amounts stated correspond to a single reaction. Holo-AcpM
592 (40 μg), in 200 mM potassium phosphate pH 7.0, 5 mM β-mercaptoethanol and a
593 final volume of 40 μL, was incubated on ice for 30 min. [2-^{14}C]Malonyl-CoA (0.05
594 μCi, 1.85 kBq, PerkinElmer) and 50 ng FabD were added and the reaction was
595 incubated at 37°C for 1 h. 42.5 μg C_{16}-Acpm:Holo-AcpM mix (42.5 μg) was added
596 with 200 mM potassium phosphate and 5 mM β-mercaptoethanol to a final volume of
597 89 μL. The reaction mix was aliquoted into 1.5 mL microcentrifuge tubes, according
598 to single assay conditions. An aliquot of KasA (0.25 μg) (or replaced with an equal
599 volume of buffer) was added and the reactions were incubated at 37°C for 1.5 h. The

600 reaction was quenched with 2 mL of freshly prepared reducing solution: 5 mg.mL^{-1}
601 $NaBH_4$ in 0.1 M K_2HPO_4, 0.4 M KCl, 30 % (v/v) tetrahydrofuran. The reaction was
602 incubated overnight at 37°C. The reduced β-ketoacyl product was extracted twice
603 using 2 mL of water-saturated toluene, and the combined organic phase washed thrice
604 using 2 mL of toluene-saturated water. The organic layer was transferred to a
605 scintillation vial and dried. The radiolabelled product was quantified by liquid
606 scintillation counting using 10 mL of EcoScintA.
607
608 **Chemoproteomics**
609
610 The chemoproteomic inhibition binding experiments were performed as previously
611 described[19,20]. Briefly, sepharose beads were derivatised with **1**, the GSK3011724A-
612 tagged analogue, at 2 mM compound concentration. Beads were washed and
613 equilibrated in lysis buffer (50 mM Tris-HCl, pH 7.4, 0.4 % Igepal-CA630, 1.5 mM
614 $MgCl_2$, 5 % glycerol, 150 mM NaCl, 25 mM NaF, 1 mM Na_3VO_4, 1 mM
615 dithiothreitol (DTT), and one Complete EDTA-free protease inhibitor tablet (Roche)).
616 The equilibrated beads were incubated at 4°C for 1 h either with 0.1 mL (0.3 mg) *M.*
617 *bovis* BCG extract or with 1 mL (5 mg) mixed HEK293/K-562/Placenta extract,
618 which was pre-incubated with compound or DMSO (vehicle control). Beads were
619 transferred either to Filter plates (Durapore (PVDF membrane, Merck Millipore)) or
620 to disposable columns (MoBiTec), washed extensively with lysis buffer and eluted
621 with SDS sample buffer. Proteins were alkylated, separated on 4–12 % Bis-Tris
622 NuPAGE (Life technologies) and stained with colloidal Coomassie. Gel lanes were
623 cut into three slices and subjected to in-gel digest using LysC for 2 h and trypsin
624 overnight[19]. Digestion, labelling with TMT isobaric mass tags, peptide fractionation,

and mass spectrometric analyses were performed[19,46]. Proteins were quantified by isobaric mass tagging and LC-MS/MS. The proteins.fasta file for *M. bovis* BCG was downloaded (May 11th 2011) from http://genome.tbdb.org/annotation/genome/tbdb/MultiDownloads.html and supplemented with the sequences of bovine serum albumin, porcine trypsin and mouse, rat, sheep and dog keratins. Decoy versions of all proteins were created and added. The search database contained a total of 11,492 protein sequences, 50 % forward, 50 % reverse. Protein identification and quantification was performed[47]. Proteins identified with >1 unique peptide matches were considered for further data analysis. Apparent dissociation constants were determined by taking into account the protein depletion by the beads[19]. Raw data tables for the chemoproteomics experiments can be found in the **Supplementary Data 1-6**.

The *M. bovis* BCG extracts were prepared as follows: *M .bovis* BCG was cultured in 7H9 medium without glycerol and supplemented with 2 % (w/v) glucose and 0.025 % (v/v) tyloxapol at 37°C for 8-10 days to reach an OD_{600nm} of 0.8-1.0. The culture was centrifuged and the pellet was washed with PBS and 0.025 % (v/v) tyloxapol. The pellet was resuspended in lysis buffer (0.4 % (v/v) Igepal was replaced with 0.8 (v/v) % NP40) and sonicated for three cycles at 50 % amplitude for 30 s (Sonics-VibracellTM) in ice. This lysate was ultracentrifuged at 4°C for 60 min and the cellular debris was discarded.

KasA protein production and structure determination

649 The over-expression plasmid, pET28a-*kasA*, was transformed into *E. coli* BL21
650 (DE3) cells. A single colony was used to inoculate 100 mL LB broth containing 50
651 $\mu g.mL^{-1}$ Kanamycin and 1 % (v/v) glucose. Cells were cultured overnight at 30°C,
652 240 rpm. A 20 L Biolaffite fermenter containing 15 L of Overnight Express Instant
653 TB (Merck), 1 % (v/v) glycerol, 50 $\mu g.mL^{-1}$ Kanamycin, and 20 mL of antifoam
654 (DC1520) was inoculated with the overnight culture to 2 % (v/v). The culture was
655 grown at 37°C, 340 rpm, 12 $L.min^{-1}$ air flow. The fermenter was cooled to 25°C
656 when the OD_{600nm} reached 2.4. The culture was incubated for 20 h before harvesting
657 the cells. The cell pellet was resuspended in 5 $mL.g^{-1}$ cells in Buffer A (50 mM Tris,
658 500 mM NaCl, 10 % (v/v) glycerol, 2 mM DTT, pH 8.5, with 1 $mg.mL^{-1}$ lysozyme,
659 Protease inhibitor cocktail set III (Sigma) and 10 µL Benzonase). The sample was
660 lysed by sonication on ice for 10 min (10 s on, 10 s off). The lysate was centrifuged
661 at 18,000 rpm, 4°C for 30 min. The supernatant was loaded onto a pre-equilibrated
662 Ni^{2+}-charged 10 mL His-Trap column. The column was washed back to baseline with
663 Buffer A and the protein was eluted using a linear gradient over 20 column volumes
664 using Buffer B (Buffer A containing 500 mM imidazole). KasA was further purified
665 by gel filtration in Buffer A using a Superdex 200 column.
666
667 KasA was co-crystallised with GSK3011724A using protein at 10.6 $mg.mL^{-1}$ and
668 ligand at a nominal concentration of 30 mM in 100 nL + 100 nL sitting drops at 20°C.
669 The well solution was 8 % (w/w) isopropanol, 0.2 M NaCl, 10 mM TCEP. Crystals
670 were cryoprotected using 30 % (v/v) glycerol prior to flash freezing in liquid nitrogen.
671 Data from a single crystal was collected at the Diamond Synchrotron Radiation
672 Facility (i04) and processed in $P3_1$ to 2.13 Å using XDS (within AUTOPROC [Global
673 Phasing Limited])[48] and AIMLESS[49]. A molecular replacement solution was

674 determined with a previously collected in house structure using Phaser[50]. The $P3_1$ cell
675 ($\alpha=\beta=90^{\circ}$, $\gamma=120^{\circ}$, a=b=77.338 Å, c=147.675 Å) has two molecules in the
676 asymmetric unit that form a dimer. Model building and refinement of the KasA
677 structures was carried out using alternating rounds of COOT[51] for manual model
678 building and REFMAC[52] for maximum likelihood refinement *via* CCP4[53]. As the data
679 was merohedrally twinned, TWIN refinement within REFMAC was used, with the
680 refined twin fraction being 39 %. A clear difference in density for GSK3011724A,
681 and also for a PEG like molecule was present in both chains in the dimer. Whilst PEG
682 was not explicitly added to the wells, this was present in adjacent crystallisation
683 conditions and we believe there may be trace PEG present. Alternatively, the linear
684 molecule may be residual lipid present in the protein, although *apo*-structures do not
685 contain this lipid and the protein is crystallised in the closed conformation even in
686 PEG conditions. A stereo diagram is displayed in **Supplementary Figure 6**. Statistics
687 for the data collection and refined co-ordinates are given in **Table 5**.
688

Data availability

The atomic coordinates and structure factors reported in this paper have been deposited in the Protein Data Bank with the code 5LD8. Data is available upon request from the corresponding authors.

References

1 Kaufmann, S. H. & van Embden, J. D. Tuberculosis: a neglected disease strikes back. *Trends Microbiol.* **1**, 2-5 (1993).

2 Zumla, A., Nahid, P. & Cole, S. T. Advances in the development of new tuberculosis drugs and treatment regimens. *Nat. Rev. Drug Discov.* **12**, 388-404 (2013).

3 World Health Organisation. Global Tuberculosis Report. *WHO* (2014).

4 Ballell, L. *et al.* Fueling open-source drug discovery: 177 small-molecule leads against tuberculosis. *ChemMedChem* **8**, 313-321 (2013).

5 Rebollo-Lopez, M. J. *et al.* Release of 50 new, drug-like compounds and their computational target predictions for open source anti-tubercular drug discovery. *PloS One* **10**, e0142293 (2015).

6 Abrahams, K. A. *et al.* Identification of novel imidazo[1,2-a]pyridine inhibitors targeting *M. tuberculosis* QcrB. *PloS One* **7**, e52951 (2012).

7 Andries, K. *et al.* A diarylquinoline drug active on the ATP synthase of *Mycobacterium tuberculosis*. *Science* **307**, 223-227 (2005).

8 Goldman, R. C. Why are membrane targets discovered by phenotypic screens and genome sequencing in *Mycobacterium tuberculosis*? *Tuberculosis* **93**, 569-588 (2013).

9 Gurcha, S. S. *et al.* Biochemical and structural characterization of mycobacterial aspartyl-tRNA synthetase AspS, a promising TB drug target. *PloS one* **9**, e113568 (2014).

10 Cox, J. A. G. *et al.* THPP target assignment reveals EchA6 as an essential fatty acid shuttle in mycobacteria. *Nat. Microbiol.* **1**, 15006 (2016).

11 Rullas, J. *et al.* Fast standardized therapeutic-efficacy assay for drug discovery against tuberculosis. *Antimicrob. Agents Chemother.* **54**, 2262-2264 (2010).

12 Zhang, M. *et al. In vitro* and *in vivo* activities of three oxazolidinones against nonreplicating *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **58**, 3217-3223 (2014).

13 Bhatt, A., Kremer, L., Dai, A. Z., Sacchettini, J. C. & Jacobs, W. R., Jr. Conditional depletion of KasA, a key enzyme of mycolic acid biosynthesis, leads to mycobacterial cell lysis. *J. Bacteriol.* **187**, 7596-7606 (2005).

14 Kremer, L. *et al.* Thiolactomycin and related analogues as novel anti-mycobacterial agents targeting KasA and KasB condensing enzymes in *Mycobacterium tuberculosis*. *J. Biol. Chem.* **275**, 16857-16864 (2000).

15 Choi, K. H., Kremer, L., Besra, G. S. & Rock, C. O. Identification and substrate specificity of beta -ketoacyl (acyl carrier protein) synthase III (mtFabH) from *Mycobacterium tuberculosis*. *J. Biol. Chem.* **275**, 28201-28207 (2000).

16 Larsen, M. H. *et al.* Overexpression of *inhA*, but not *kasA*, confers resistance to isoniazid and ethionamide in *Mycobacterium smegmatis*, *M. bovis* BCG and *M. tuberculosis*. *Mol. Microbiol.* **46**, 453-466 (2002).

17 Coxon, G. D. *et al.* Synthesis, antitubercular activity and mechanism of resistance of highly effective thiacetazone analogues. *PloS One* **8**, e53162 (2013).

18 Kremer, L. *et al.* Mycolic acid biosynthesis and enzymic characterization of the beta-ketoacyl-ACP synthase A-condensing enzyme from *Mycobacterium tuberculosis*. *Biochem. J.* **364**, 423-430 (2002).

745 19 Bantscheff, M. *et al.* Chemoproteomics profiling of HDAC inhibitors reveals
746 selective targeting of HDAC complexes. *Nat. Biotechnol.* **29**, 255-265 (2011).
747 20 Savitski, M. M. *et al.* Targeted data acquisition for improved reproducibility
748 and robustness of proteomic mass spectrometry assays. *J. Am. Soc. Mass*
749 *Spectrom.* **21**, 1668-1679 (2010).
750 21 Sirakova, T. D., Dubey, V. S., Cynamon, M. H. & Kolattukudy, P. E.
751 Attenuation of *Mycobacterium tuberculosis* by disruption of a mas-like gene
752 or a chalcone synthase-like gene, which causes deficiency in dimycocerosyl
753 phthiocerol synthesis. *J. Bacteriol.* **185**, 2999-3008 (2003).
754 22 Waddell, S. J. *et al.* Inactivation of polyketide synthase and related genes
755 results in the loss of complex lipids in *Mycobacterium tuberculosis* H37Rv.
756 *Lett. App. Microbiol.* **40**, 201-206 (2005).
757 23 Monetti, M., Nagaraj, N., Sharma, K. & Mann, M. Large-scale phosphosite
758 quantification in tissues by a spike-in SILAC method. *Nat. Methods* **8**, 655-
759 658 (2011).
760 24 Siragusa, L., Cross, S., Baroni, M., Goracci, L. & Cruciani, G. BioGPS:
761 navigating biological space to predict polypharmacology, off-targeting, and
762 selectivity. *Proteins* **83**, 517-532 (2015).
763 25 Schiebel, J. *et al.* Structural basis for the recognition of mycolic acid
764 precursors by KasA, a condensing enzyme and drug target from
765 *Mycobacterium tuberculosis*. *J. Biological Chem.* **288**, 34190-34204 (2013).
766 26 Luckner, S. R., Machutta, C. A., Tonge, P. J. & Kisker, C. Crystal structures
767 of *Mycobacterium tuberculosis* KasA show mode of action within cell wall
768 biosynthesis and its inhibition by thiolactomycin. *Structure* **17**, 1004-1013
769 (2009).

27 Maitra, A. *et al.* Repurposing-a ray of hope in tackling extensively drug resistance in tuberculosis. *Int. J. Infect. Dis.* **32**, 50-55 (2015).

28 Harbut, M. B. *et al.* Auranofin exerts broad-spectrum bactericidal activities by targeting thiol-redox homeostasis. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 4453-4458 (2015).

29 Kana, B. D., Karakousis, P. C., Parish, T. & Dick, T. Future target-based drug discovery for tuberculosis? *Tuberculosis* **94**, 551-556 (2014).

30 Moreira, W. *et al.* Target mechanism-based whole-cell screening identifies bortezomib as an inhibitor of caseinolytic protease in mycobacteria. *MBio.* **6**, e00253-00215 (2015).

31 Matei, L. *et al.* Synthesis and biological activities of some new isonicotinic acid 2-(2-hydroxy-8-substituted-tricyclo[7.3.1.0(2.7)]tridec-13-ylidene)-hydrazides. *Bioorg. Med. Chem.* **23**, 401-410 (2015).

32 Nigam, A. *et al.* Modification of rifamycin polyketide backbone leads to improved drug activity against rifampicin-resistant *Mycobacterium tuberculosis*. *J. Biol. Chem.* **289**, 21142-21152 (2014).

33 Bhatt, A., Molle, V., Besra, G. S., Jacobs, W. R., Jr. & Kremer, L. The *Mycobacterium tuberculosis* FAS-II condensing enzymes: their role in mycolic acid biosynthesis, acid-fastness, pathogenesis and in future drug development. *Mol. Microbiol.* **64**, 1442-1454 (2007).

34 Marrakchi, H. *et al.* MabA (FabG1), a *Mycobacterium tuberculosis* protein involved in the long-chain fatty acid elongation system FAS-II. *Microbiol.* **148**, 951-960 (2002).

793 35 Sacco, E. *et al.* The missing piece of the type II fatty acid synthase system
794 from *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 14628-
795 14633 (2007).
796 36 Banerjee, A. *et al. inhA*, a gene encoding a target for isoniazid and
797 ethionamide in *Mycobacterium tuberculosis*. *Science* **263**, 227-230 (1994).
798 37 Gande, R. *et al.* Acyl-CoA carboxylases (*accD2* and *accD3*), together with a
799 unique polyketide synthase (Cg-pks), are key to mycolic acid biosynthesis in
800 Corynebacterianeae such as *Corynebacterium glutamicum* and *Mycobacterium*
801 *tuberculosis*. *J. Biol. Chem.* **279**, 44847-44857 (2004).
802 38 Portevin, D. *et al.* A polyketide synthase catalyzes the last condensation step
803 of mycolic acid biosynthesis in mycobacteria and related organisms. *Proc.*
804 *Natl. Acad. Sci. U. S. A.* **101**, 314-319 (2004).
805 39 Bhatt, A., Brown, A. K., Singh, A., Minnikin, D. E. & Besra, G. S. Loss of a
806 mycobacterial gene encoding a reductase leads to an altered cell wall
807 containing beta-oxo-mycolic acid analogs and accumulation of ketones. *Chem.*
808 *Biol.* **15**, 930-939 (2008).
809 40 Parrish, N. M., Kuhajda, F. P., Heine, H. S., Bishai, W. R. & Dick, J. D.
810 Antimycobacterial activity of cerulenin and its effects on lipid biosynthesis. *J.*
811 *Antimicrob. Chemother.* **43**, 219-226 (1999).
812 41 Brown, A. K., Taylor, R. C., Bhatt, A., Futterer, K. & Besra, G. S.
813 Platensimycin activity against mycobacterial beta-ketoacyl-ACP synthases.
814 *PloS One* **4**, e6306 (2009).
815 42 Remuinan, M. J. *et al.* Tetrahydropyrazolo[1,5-a]pyrimidine-3-carboxamide
816 and N-benzyl-6',7'-dihydrospiro[piperidine-4,4'-thieno[3,2-c]pyran] analogues

with bactericidal efficacy against *Mycobacterium tuberculosis* targeting MmpL3. *PloS One* **8**, e60933 (2013).

43 Collins, L. & Franzblau, S. G. Microplate alamar blue assay versus BACTEC 460 system for high-throughput screening of compounds against *Mycobacterium tuberculosis* and *Mycobacterium avium*. *Antimicrob. Agents Chemother.* **41**, 1004-1009 (1997).

44 Besra, G. S. Preparation of cell-wall fractions from mycobacteria. In *Methods in Molecular Biology: Mycobacteria Protocols*. Editors Parish, T. & Stoker, N. G. New Jersey: Humana Press. **101**, 91-107 (1998).

45 Kremer, L. *et al.* Biochemical characterization of acyl carrier protein (AcpM) and malonyl-CoA:AcpM transacylase (mtFabD), two major components of *Mycobacterium tuberculosis* fatty acid synthase II. *J. Biological. Chem.* **276**, 27967-27974 (2001).

46 Werner, T. *et al.* Ion coalescence of neutron encoded TMT 10-plex reporter ions. *Anal. Chem.* **86**, 3594-3601 (2014).

47 Savitski, M. M. *et al.* Measuring and managing ratio compression for accurate iTRAQ/TMT quantification. *J. Proteome Res.* **12**, 3586-3598 (2013).

48 Kabsch, W. Xds. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125-132 (2010).

49 Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D. Biol. Crystallogr.* **69**, 1204-1214 (2013).

50 McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674 (2007).

51 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126-2132 (2004).

52 Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D. Biol. Crystallogr.* **53**, 240-255 (1997).

53 Collaborative Computational Project, N. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **50**, 760-763 (1994).

54 Sridharan, S. *et al.* X-ray crystal structure of *Mycobacterium tuberculosis* beta-ketoacyl acyl carrier protein synthase II (mtKasB). *J. Mol. Biol.* **366**, 469-480 (2007).

55 Sachdeva, S. *et al.* Separate entrance and exit portals for ligand traffic in *Mycobacterium tuberculosis* FabH. *Chem. Biol.* **15**, 402-412 (2008).

End Notes

Acknowledgements

GSB acknowledges support in the form of a Personal Research Chair from Mr James Bardrick and a Royal Society Wolfson Research Merit Award. The research leading to these results has received funding from the European Union´s 7th framework programme (FP7- 2007-2013) under Grant Agreement No 261378 and the Medical Research Council (MR/K012118/1). We thank Nico Zinn and Toby Mathieson for mass spectrometry and database design and David Dixon for help with the SPR. We also thank the TB Alliance for their helpful discussions and expertise in the field.

Author contributions

Conceived and designed the experiments: KAA, SGD, CC, JR, MJRL, JL, MB, IAB, MCI, NC, AM, MSMM, CS, GD, LB, DB, RHB, GSB. Performed the experiments: KAA, SGD, JAGC, MN, AS, AA, PH, CS, MA, SSG, RC, CC, SGD, LK, MJRL, JR, RHB. Analysed the data: KAA, CC, JR, MJRL, MB, NJL, SGD, MSMM, EJN, AM, MCI, NC, DB, LB, GSB, RHB. Wrote the paper: KAA, LB, SGD, GD, JAGC, PJM, CC, GSB, RHB.

In memoriam statement

In memoriam of Emilio Alvarez Ruiz, an example of professionalism, comradeship and inspiration. His scientific knowledge, advice and wisdom will be greatly missed.

878

879 **Competing financial interests**: The authors declare no competing financial interests.

880 **Figure Legends**

881

882 **Figure 1. The anti-tubercular activity of GSK3011724A.** (**a**) Chemical structure of
883 GSK3011724A. (**b, c**) GSK3011724A was evaluated in an acute (**b**) and chronic (**c**)
884 model of murine TB infection. Each symbol represents data from a single mouse. A
885 two-fold log CFU reduction equates to a 2 log CFU difference compared to untreated
886 mice at the end of the treatment period. Untreated mice gave a log CFU count of 7.1 ±
887 0.1 (mean ± SD, *n*=5 mice) in the acute assay and 5.8 ± 0.26 (mean ± SD, *n*= 7 mice
888 for the INH experiment) or 6.0 ± 0.5 (mean ± SD, *n*= 5 mice for GSK3011724A
889 experiment) in the chronic assay. (**d**) Comparison of the *in vivo* efficacy of
890 GSK3011724A and INH against *M. tuberculosis* H37RV wild-type (H37RV WT) and
891 *M. tuberculosis* KasA mutant strains. Mice were infected with each strain and
892 administered with INH or GSK3011724A at different doses during the acute phase
893 (days 1 to 8 after infection). Each dot represents data from a single mouse. Log CFU
894 counts are shown as the difference with respect to the untreated group infected with
895 each strain (ΔlogCFU for each mouse).

896

897

898 **Figure 2. Inhibition of mycolic acid biosynthesis by GSK3011724A.** (**a**) *M. bovis*
899 BCG cultures, labelled with [^{14}C]-acetate, were treated with GSK3011724A, TLM
900 and INH. The total FAMEs and MAMEs were extracted and analysed by
901 autoradiography-TLC using equal counts (25,000 cpm) for each lane, respectively
902 (left and middle panels). In addition, cell wall bound MAMEs were isolated and an
903 equal aliquot (5 %) was analysed by autoradiography-TLC for each lane, respectively
904 (right panel). (**b**) *M. bovis* BCG total cellular lipids containing TDM/TMM were
905 extracted following labelling and drug treatment (GSK3011724A, TLM and INH) and
906 analysed by autoradiography-TLC using equal counts (25,000 cpm) for each lane,
907 respectively. (**c**) Impact on the MIC of GSK3011724A upon the over-expression of
908 members of FAS-II in *M. bovis* BCG. The over-expression constructs of *M.*
909 *tuberculosis* enzymes (Mt) pMV261-Mt-*kasA*, pMV261-Mt-*kasB*, pMV261-Mt-*fabH*,
910 pMV261-Mt-*inhA*, pMV261-Mt-*hadABC* and pMV261-Mt-*mabA* were electroporated
911 into *M. bovis* BCG and the MIC of GSK3011724A was evaluated with reference to
912 *M. bovis* BCG pMV261. The outer and inner spots represent 10^3 and 10^2 cells plated,
913 respectively.
914
915
916

Figure 3. Chemoproteomics profiling of GSK3011724A. (**a**) A propylamine-tagged derivative of GSK3011724A (**1**, inset) was synthesised and covalently immobilised to NHS-activated sepharose. Beads were incubated with *M. bovis* BCG extract either in the presence of vehicle (DMSO) or GSK3011724A (10 μM, 40 μM). Proteins captured by the beads in both conditions were quantified by LC-MS/MS analysis. KasA, Pks10 and Pks11 were identified as potential targets of GSK3011724A by virtue of their reduced capturing in the presence of excess GSK3011724A. (**b**) Generation of IC_{50} values for KasA, Pks10 and Pks11. The chemoproteomic experiment was performed as in ***a*** but over a range of concentrations of the competing 'free' inhibitor GSK3011724A (2 μM – 0.003 μM for KasA, 40 μM – 0.16 μM for Pks10 and Pks11) and a structurally related inactive analogue, **2** (40 μM – 0.16 μM). Apparent dissociation constants for GSK3011724A were determined from two independent experiments.

Figure 4. X-ray crystallographic analysis of KasA complexed with GSK3011724A. (**a**) Co-crystal structure of KasA dimer (yellow, blue) with GSK3011724A (green space filled) and PEG (green sticks) bound with the open conformation of the acyl channel. The active site cysteine C171 is shown in line format, as are an overlay of the TLM and a phospholipid ligand taken from PDB entry 4C72[25]. (**b**) GSK3011724A and PEG both shown in stick format filling the acyl channel. (**c**) Hydrogen bonds between GSK3011724A and E199 within the binding site.

Figure 5. Impact of resistance-conferring mutations on the KasA-GSK3011724A crystal structure and structural comparison of KasA with KasB and FabH. (**a**) Mapping of residues where resistance mutations occur onto the dimeric KasA-GSK3011724A X-ray crystallographic complex. GSK3011724A (green stick) marks the acyl channel, whereas TLM (magenta) from PDB entry 4C72[25] marks the malonyl pocket. (**b**) Residues P201 and P206 that confer resistance by direct interactions with GSK3011724A are shown, as are the proximity of other resistance positions visible. (**c**) A ribbon representation of the open and closed forms of KasA with resistance mutations highlighted. The open KasA-GSK3011724A dimer is shown in yellow and blue (chain A, B respectively). The apo, closed conformation (PDB entry 2WGD[26]) is shown in orange and purple. The HTH arm formed by $\alpha 5$-$\alpha 6$ and the $\alpha' 2$ from its dimer partner move substantially on transition between the open and closed states. Residues within these two regions of movement where resistance mutations are found are shown in stick format. GSK3011724A, PEG and other residues sensitive to resistance mutations are shown in line format**.** (**d**) Surface around Chain A within KasA-GSK3011724A complex shown to illustrate the role of M277 in the open and closed states. Colouring and representations identical to ***c***. (**e**) Hydrogen bonding interactions around T114 within the open KasA-GSK3011724A complex. (**f**) Structural alignment of KasA-GSK3011724A (yellow and blue) with KasB (PDB entry 2GP6[54]) (magenta). The indazole ring of GSK3011724A sits favourably over G200 and P201 in KasA. In KasB, the respective residues of T199 and R200 would not be able to accommodate this favourable binding mode. (**g**) Structural overlay of KasA-GSK3011724A (yellow and blue) with FabH (PDB entry 2QNZ[55]) (orange). In FabH, T145 would directly clash with the indazole ring of GSK3011724A, accounting for the compound selectivity for KasA.

966 **Tables**

967

968

969

970

971

Assay	Value
MIC H37Rv	0.8 μM
HepG2 IC_{50}	>100 μM
$ChromLogD_{7.4}$	3.8
Kinetic Solubility	>550 μM
hERG IC_{50}	>50 μM
CYP3A4 IC_{50}	>40 μM
CYP1A2 IC_{50}	>50 μM
CYP2C9 IC_{50}	12.6 μM
CYP2D6 IC_{50}	>50 μM
18 Gram-positive and Gram-negative species	>128 $\mu g.mL^{-1}$
Plasma Protein Binding (mouse)	76.9%
Plasma Protein Binding (human)	91.8%
Cli (mouse)	6.1 $mL.min^{-1}g^{-1}$
Cli (human)	0.95 $mL.min^{-1}g^{-1}$

972

973 **Table 1. *In vitro* profile of GSK3011724A.**

974

975

976

977

Target Dose ($mg.kg^{-1}$)	C_{max} ($ng.mL^{-1}$)	C_{max}/Dose $ng.mL^{-1}$ ($mg.kg^{-1}$)	t_{max} (h)	AUC_{0-t} ($ng.h.mL^{-1}$)	DNAUC $ng.h.mL^{-1}$ ($mg.kg^{-1}$)
10[a]	917±279	84.7±21.5	0.25-0.75	1099±152	104±19.1
25	1392±424	63.9±19.4	0.5	2279±581	107±27.3
50	12500±153	250±23.1	0.25-1.0	32549±3009	667±64.2
100	10570±572	120±6.49	0.75-1.0	57179±7559	667±57.2
200	33000±7017	165±35.1	2.0	251437±37054	1272±188

[a]Mean and SD values calculated from *n* = 3 mice.

Table 2. *In vivo* pharmacokinetic profile of GSK3011724A. Pharmacokinetic parameters estimated for GSK3011724A after a single oral gavage administration to female C57BL/6J mice (*n* = 3) at different doses. Mice dosed above 400 $mg.kg^{-1}$ were withdrawn after the second administration of GSK3011724A due to poor clinical status. Standard Deviation (SD) of the mean is also included. C_{max} = maximum concentration measured; AUC = Area under the curve; DNAUC = Dose normalised area under the curve.

Organism	Mutant	Frequency of SNP	Genome position of SNP	Base change	Amino acid substitution
M. bovis BCG	1	75 %	2497100	aCc/aGc	T114S
		2.44 %	2497376	cCc/cTc	P206L
	2	100 %	2497142	cTg/cCg	L128P
	3	100 %	2497373	cTg/cCg	L205P
	4	100 %	2497376	cCc/cTc	P206L
	5	100 %	2497195	Atg/Gtg	M146V
	6	100 %	2497589	aTg/aCg	M277T
M. tuberculosis	1	-	2518524	gTg/gCg	V137A
	2	-	2518524	gTg/gYg	V137A/V
	3	-	2518524	gTg/gCg	V137A
	4	-	2518524	gTg/gCg	V137A
	5	-	2518715	Ccc/Tcc	P201S
	6	-	2518715	Ccc/Tcc	P201S
	7	-	2518715	Ccc/Tcc	P201S
	8	-	2518715	Ccc/Tcc	P201S
	9	-	2518943	Atg/Gtg	M277V
	10	-	2518944	aTg/aCg	M277T
	11	-	2519327	Ggc/Agc	G405S

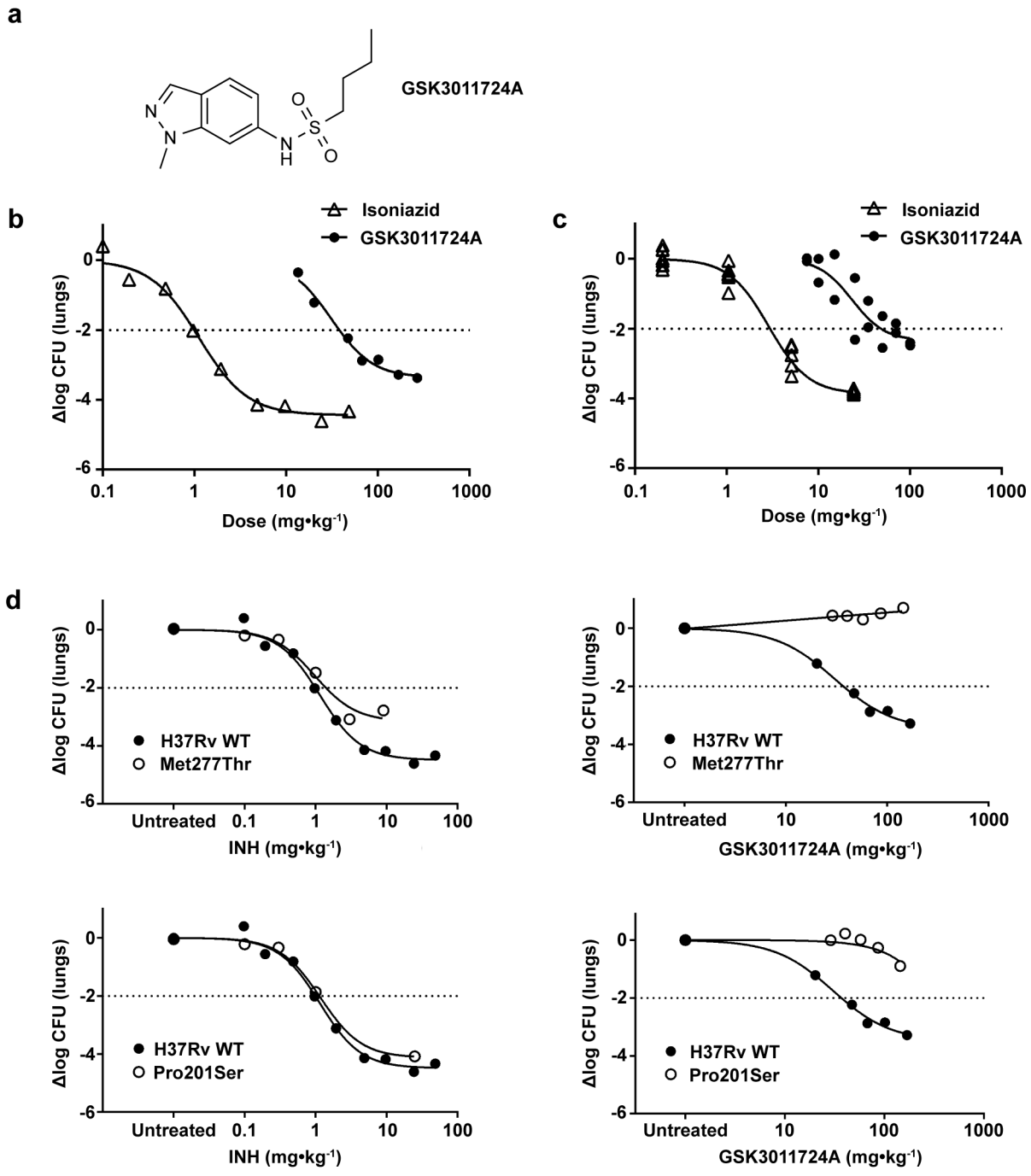
Table 3: Sequencing of spontaneous resistant mutants identifies SNPs in *kasA*. WGS and sequencing of *kasA* from GSK3011724A-resistant mutants of *M. bovis* BCG and *M. tuberculosis,* respectively, revealed *kasA*-specific mutations. For the WGS of *M. bovis* BCG mutants, the frequency of SNP is shown. This corresponds to the percentage of SNP in the cell population from which the genomic DNA was extracted. The *kasA* SNP frequency in the *M. tuberculosis* mutants cannot be deduced due to the method of analysis used. The genomic location of the mutations is stated. The missense mutations are represented in codons, where the capital letter identifies the base change. The 'Y' indicates a mixed population of 'C' or 'T' bases. The resulting amino acid substitutions are listed.

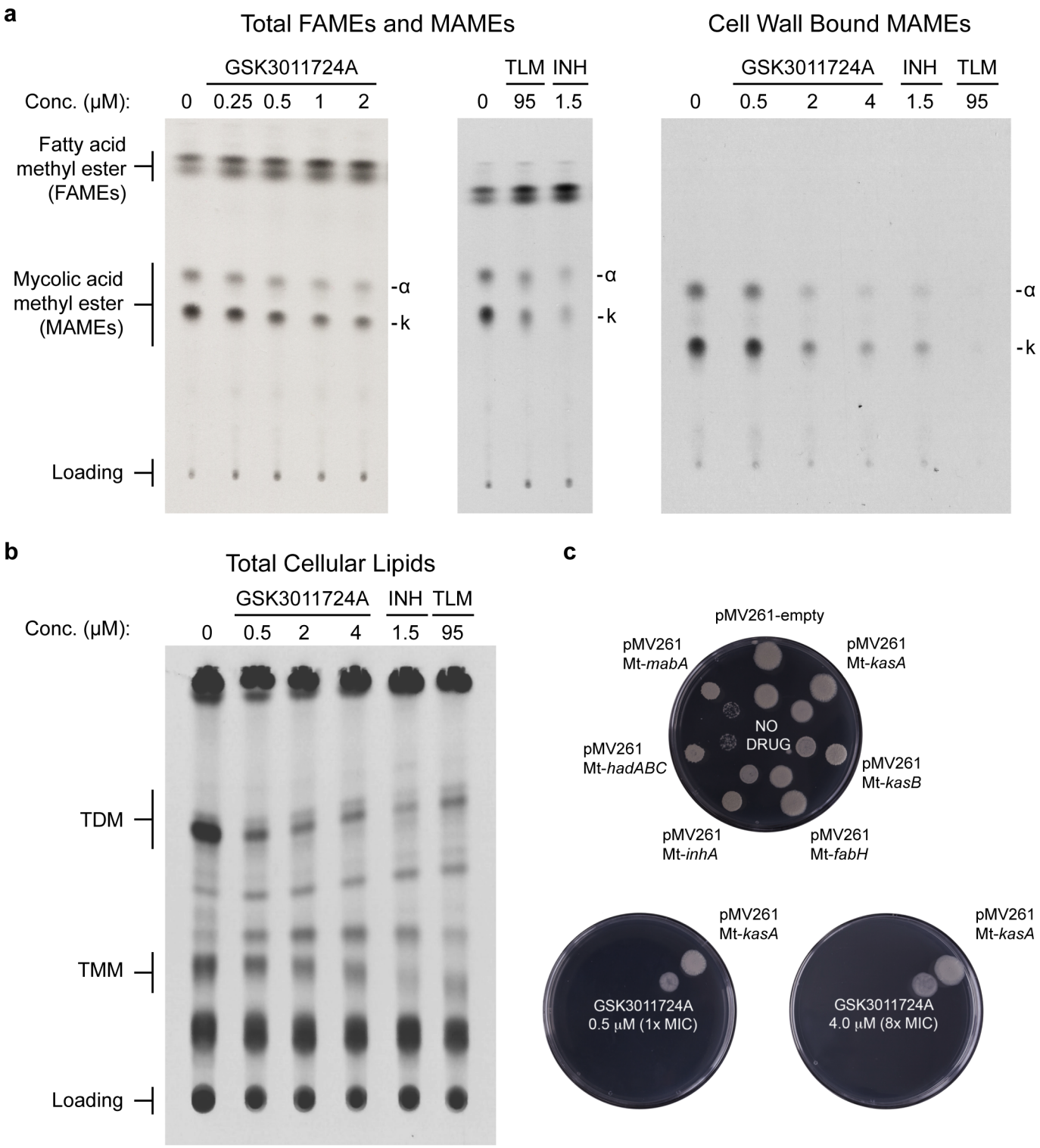
Inhibitor	Inhibitor concentration (μM)	FAMEs		α-MAMEs		k-MAMEs		TDM		TMM	
		cpm	% relative to control	cpm	% relative to control	cpm	% relative to control	cpm	% relative to control	cpm	% relative to control
Control	0	9796	100	4029	100	7641	100	3149	100	1425	100
	0.25	16207	165	3438	85	7147	94	-[ND]	-[ND]	-[ND]	-[ND]
	0.5	17780	182	3026	75	6651	87	864	27	911	64
GSK3011724A	1	16478	168	2488	62	4925	64	874	28	893	63
	2	18268	186	1806	45	4407	58	-[ND]	-[ND]	-[ND]	-[ND]
	4	-[ND]	-[ND]	-[ND]	-[ND]	-[ND]	-[ND]	652	21	951	67
Control	0	10231	100	4156	100	6722	100	-[*]	-[*]	-[*]	-[*]
TLM	95	15334	150	2721	65	3727	55	817	26	704	49
INH	1.5	16926	165	1426	34	1808	27	642	20	552	39

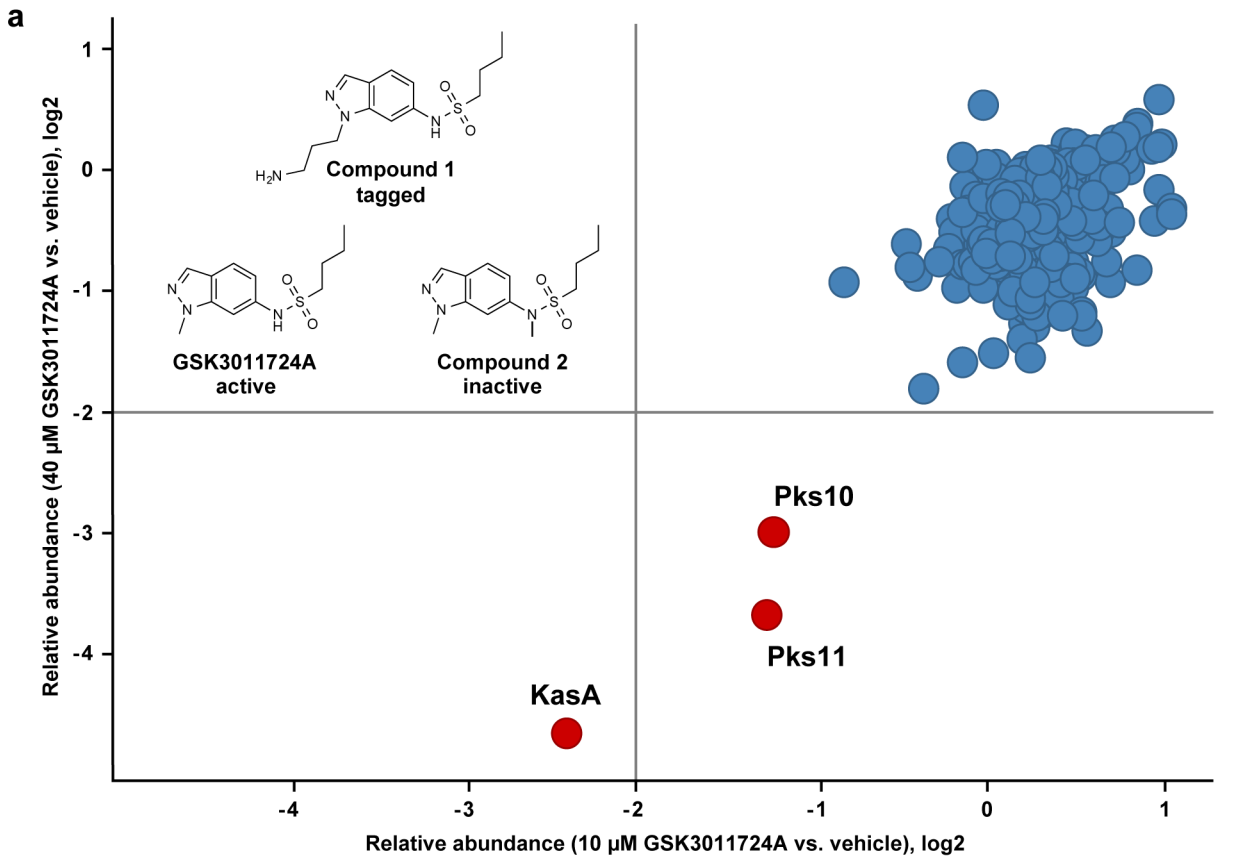
Table 4: Quantification of inhibition of mycolate containing products by GSK3011724A. Quantification of [^{14}C]-labelled products were determined by excising spots directly from TLC plates (**Fig. 2a**, left and middle panels, and **Fig. 2b**) and subjecting them to scintillation counting using 10 mL of EcoScintA. *Control values as reported above; ND, not determined.

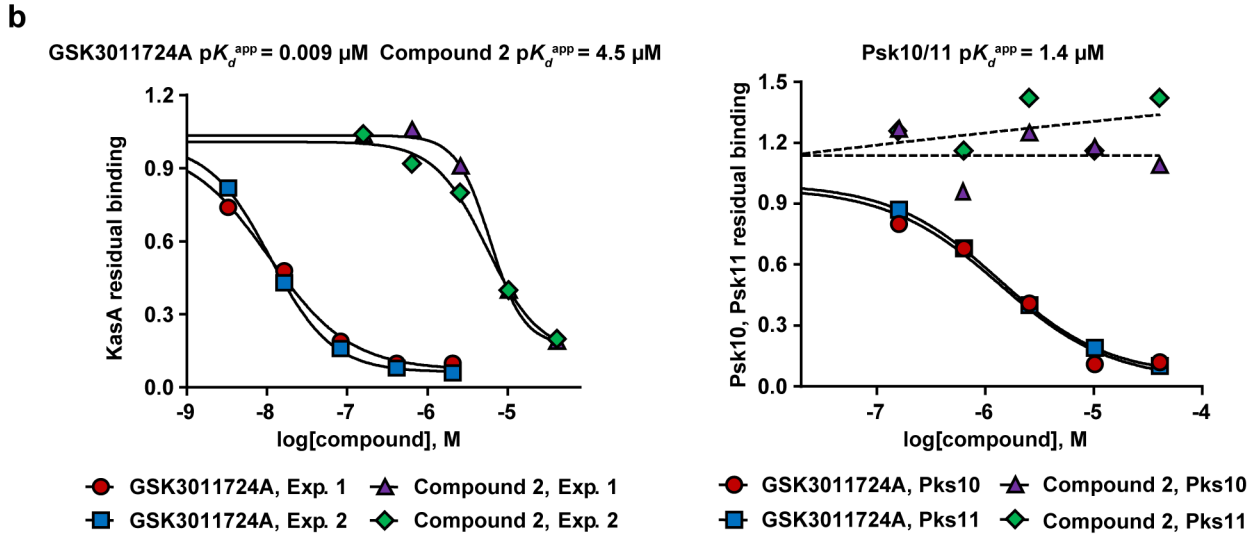
X-ray diffraction data	KasA / GSK3011724A
Data collection	
Space group	$P3_1$
Cell dimensions	
a, *b*, *c* (Å)	77.34, 77.34, 147.68
α, β, γ (°)	90.000, 90.000, 120.000
Resolution (Å)	66.98-2.13(2.38-2.13)
R_{merge}	0.053 (0.474)
ccI1/2	0.998 (0.822)
$I/\sigma I$	16.9 (3.0)
Completeness (%)	98.6 (97.3)
Redundancy	4.3(4.0)
Wilson B-factor	34.16
Refinement	
Resolution (Å)	66.98-2.13
No. reflections	235866 (60834)
No. uniq reflections	54699 (15385)
$R_{work/}$ R_{free}	0.162/0.186
No. atoms	6470
Protein	6041
Ligand/ion	36/28
Water	365
B-factors	
Protein	39.70
Ligand/ion	39.90/48.61
Water	43.92
R.m.s deviations	
Bond lengths (Å)	0.0044
Bond angles (°)	0.8045
Twin fraction	0.39

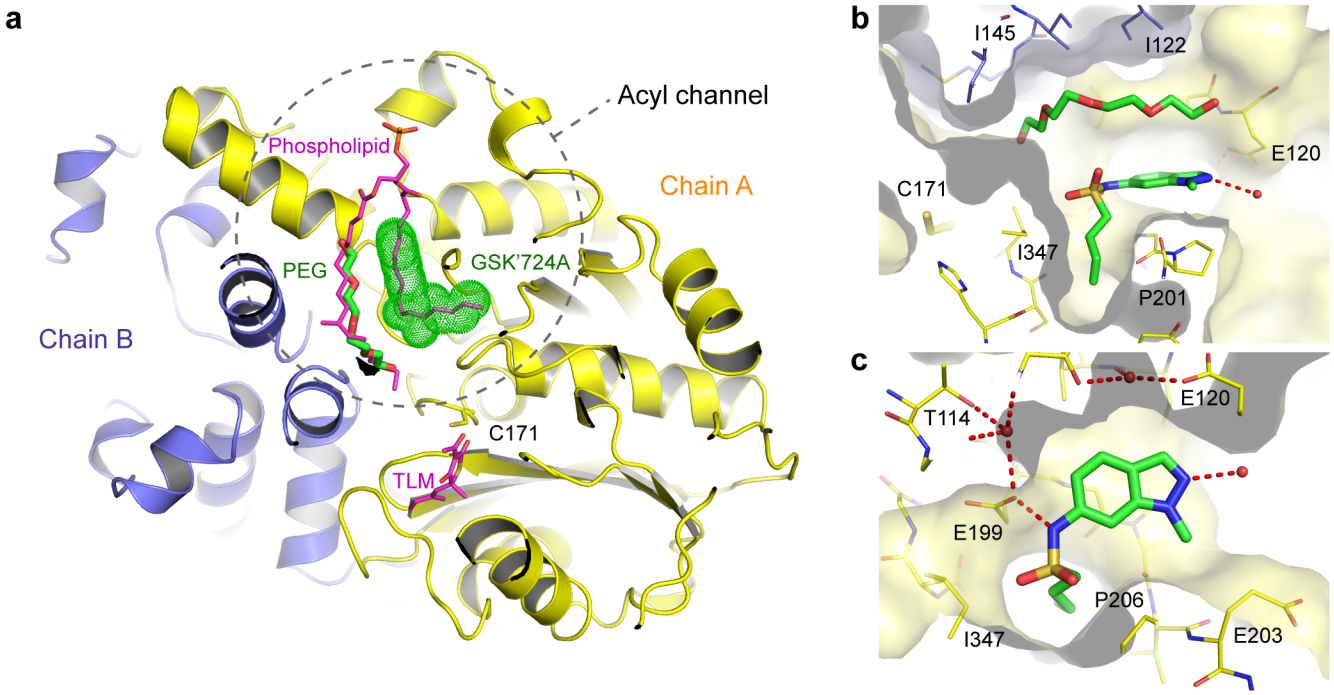
Table 5: Data collection and refinement statistics. Values in parentheses are for the highest-resolution shell. Collection based on a single crystal.

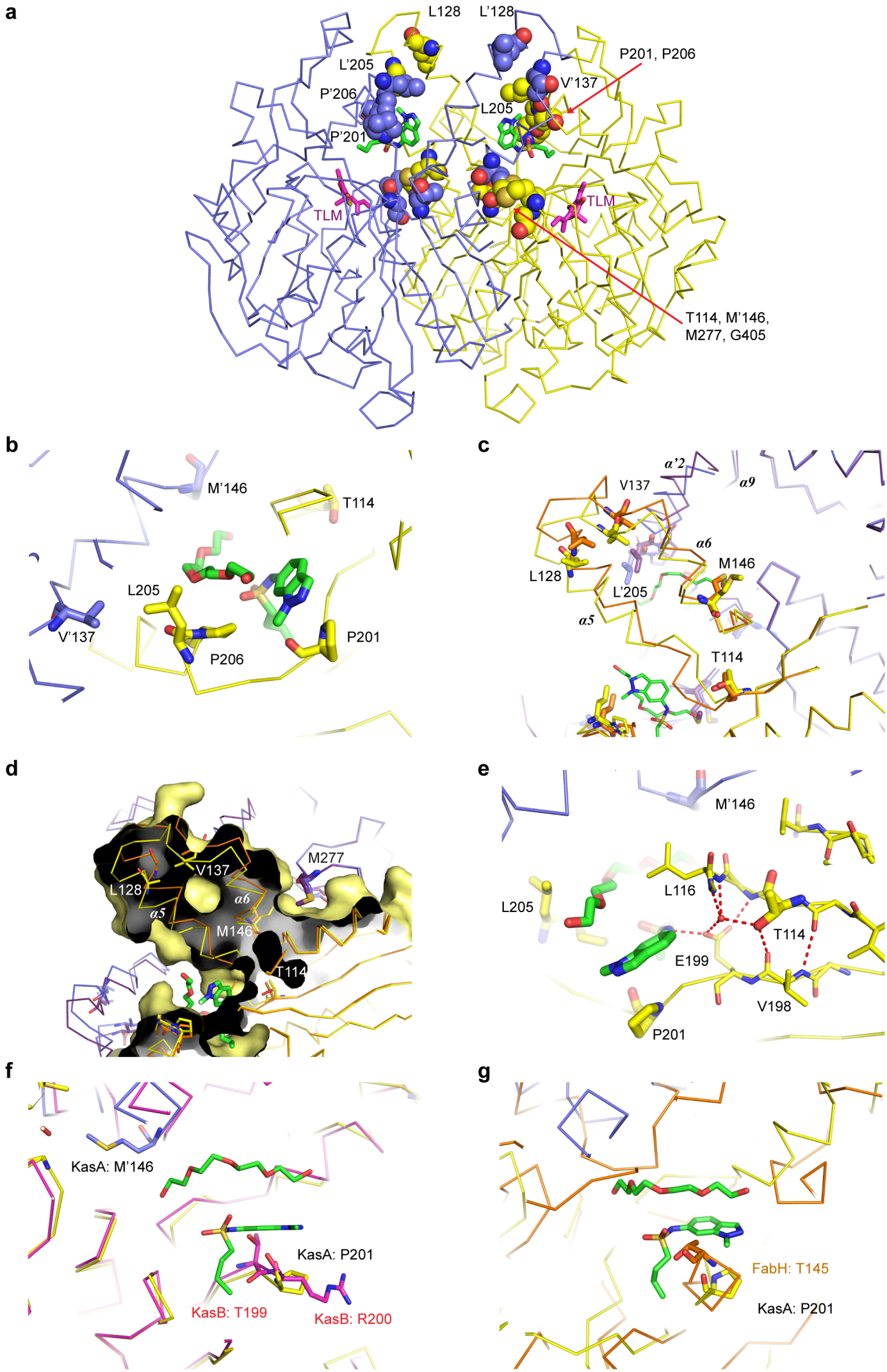

a
GSK3011724A
b
Isoniazid
GSK3011724A
Δlog CFU (lungs)
Dose (mg•kg-1)
c
Isoniazid
GSK3011724A
Δlog CFU (lungs)
Dose (mg•kg-1)
d
H37Rv WT
Met277Thr
Δlog CFU (lungs)
Untreated
INH (mg•kg-1)
H37Rv WT
Met277Thr
Δlog CFU (lungs)
GSK3011724A (mg•kg-1)
H37Rv WT
Pro201Ser
Δlog CFU (lungs)
Untreated
INH (mg•kg-1)
H37Rv WT
Pro201Ser
Δlog CFU (lungs)
GSK3011724A (mg•kg-1)


a
Total FAMEs and MAMEs
Cell Wall Bound MAMEs
GSK3011724A
TLM
INH
Conc. (μM):
0
0.25
0.5
1
2
95
1.5
4
Fatty acid methyl ester (FAMEs)
Mycolic acid methyl ester (MAMEs)
-α
-k
Loading
b
Total Cellular Lipids
TDM
TMM
c
pMV261-empty
pMV261 Mt-mabA
pMV261 Mt-kasA
pMV261 Mt-hadABC
pMV261 Mt-kasB
pMV261 Mt-inhA
pMV261 Mt-fabH
NO DRUG
GSK3011724A 0.5 μM (1x MIC)
GSK3011724A 4.0 μM (8x MIC)


a
Relative abundance (40 µM GSK3011724A vs. vehicle), log2
Relative abundance (10 µM GSK3011724A vs. vehicle), log2
Compound 1
tagged
GSK3011724A
active
Compound 2
inactive
Pks10
Pks11
KasA

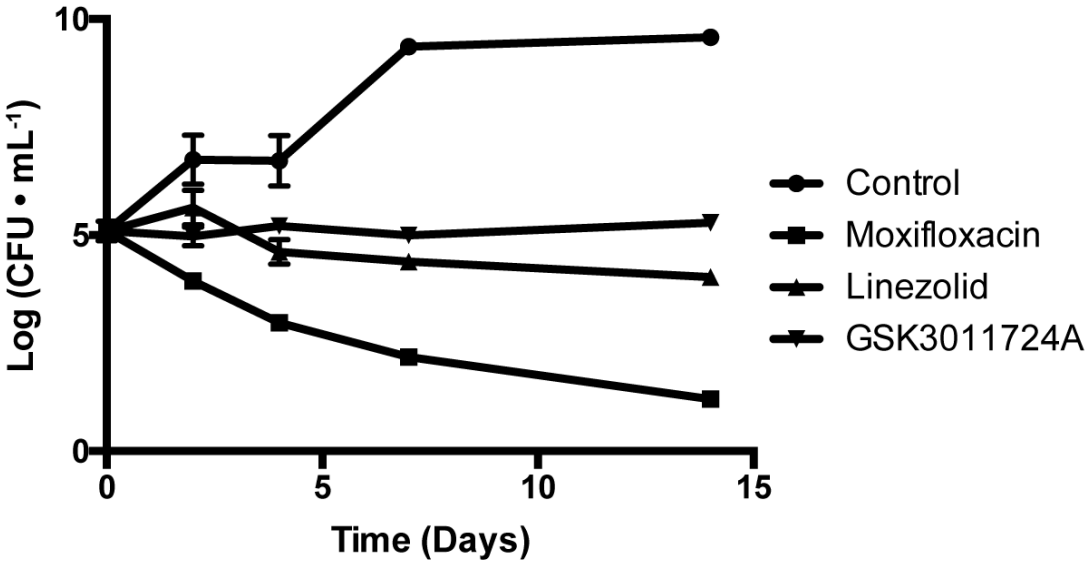

b
GSK3011724A pK_d^{app} = 0.009 µM Compound 2 pK_d^{app} = 4.5 µM
KasA residual binding
log[compound], M
GSK3011724A, Exp. 1
GSK3011724A, Exp. 2
Compound 2, Exp. 1
Compound 2, Exp. 2
Psk10/11 pK_d^{app} = 1.4 µM
Psk10, Psk11 residual binding
log[compound], M
GSK3011724A, Pks10
GSK3011724A, Pks11
Compound 2, Pks10
Compound 2, Pks11


a
Acyl channel
Phospholipid
Chain A
PEG
GSK'724A
Chain B
C171
TLM
b
I145
I122
E120
C171
I347
P201
c
E120
T114
E199
P206
I347
E203

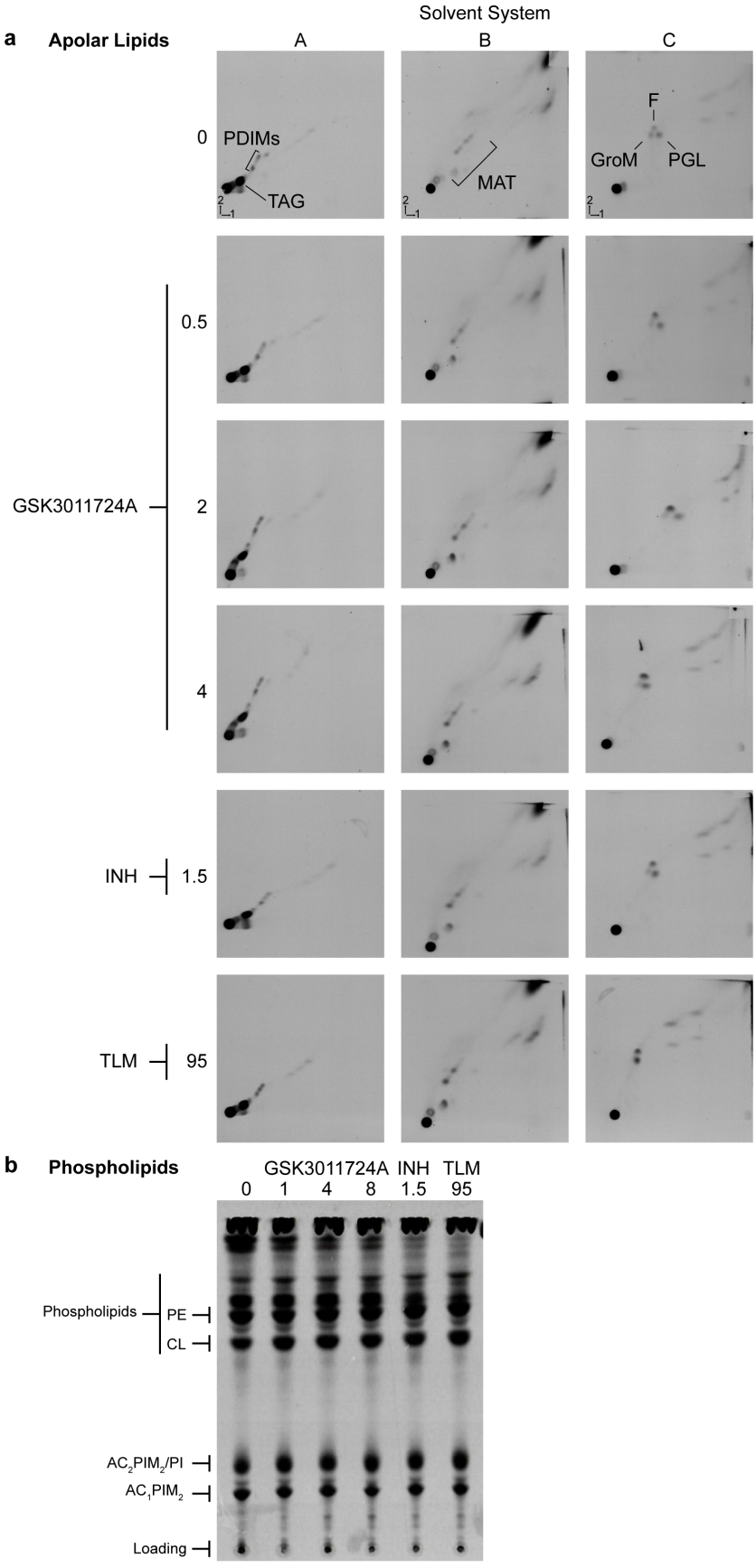

a
L128
L'128
L'205
P'206
P'201
L205
V'137
P201, P206
TLM
TLM
T114, M'146,
M277, G405
b
M'146
T114
L205
V'137
P206
P201
c
α'2
α9
V137
α6
M146
L128
L'205
α5
T114
d
M277
V137
L128
α6
α5
M146
T114
e
M'146
L116
L205
T114
E199
V198
P201
f
KasA: M'146
KasA: P201
KasB: T199
KasB: R200
g
FabH: T145
KasA: P201

Supplementary Information

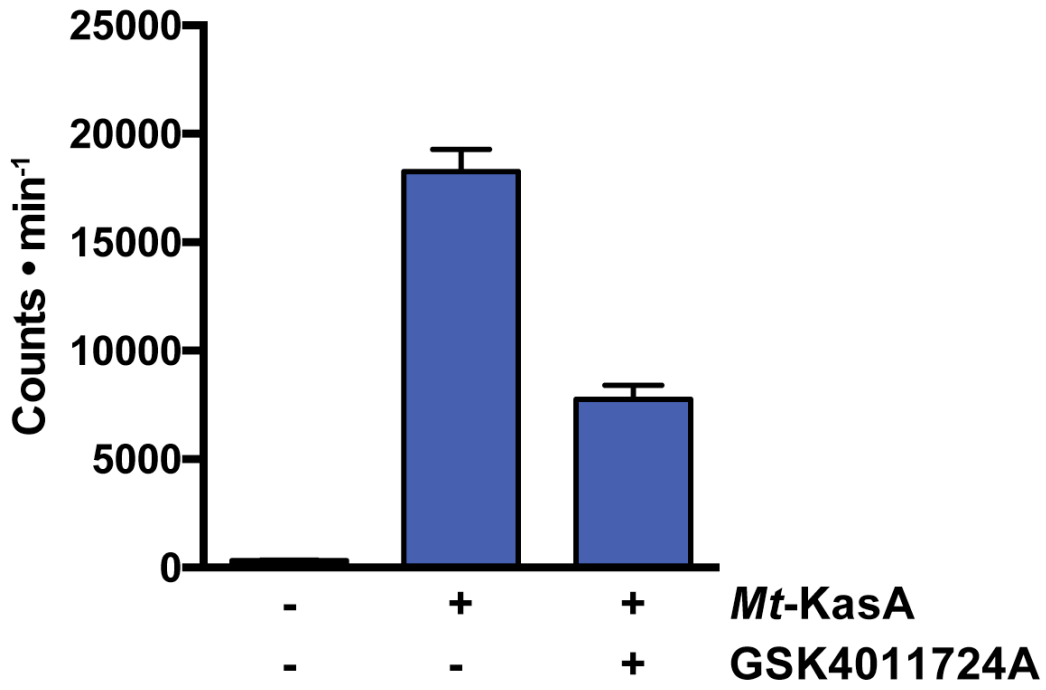
Supplementary Figures



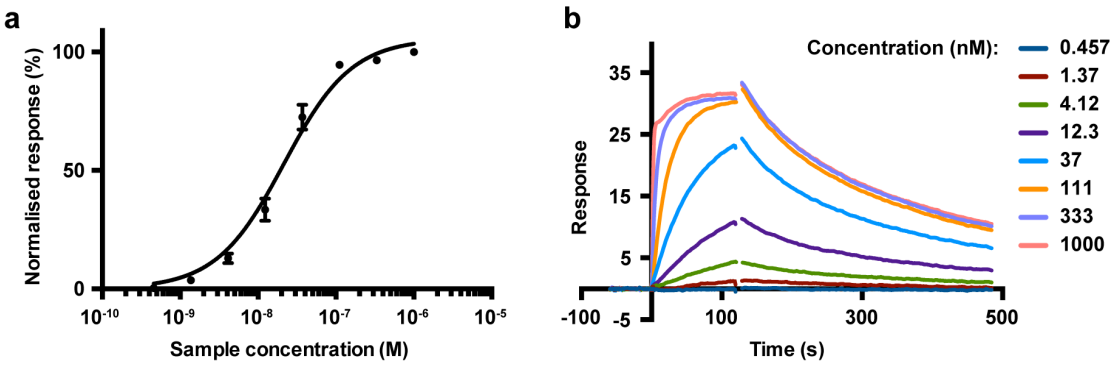
Supplementary Figure 1: *In vitro* death curve analysis of GSK3011724A. *M. tuberculosis* was treated with GSK3011724A at 0.5 μM (10 x MIC in liquid media), along with the bacteriostatic and bactericidal compounds, linezolid (17.8 μM) and moxifloxacin (1.5 μM), respectively (10 x MIC), and compared to untreated cells (control). CFU counts were performed in triplicate and analysed and displayed as mean ± standard deviation. GSK3011724A exhibited bacteriostatic activity in this *in vitro* system, despite its bactericidal activity *in vivo*.


a
Apolar Lipids
Solvent System
A
B
C
0
PDIMs
TAG
MAT
F
GroM
PGL
GSK3011724A
0.5
2
4
INH
1.5
TLM
95
b
Phospholipids
GSK3011724A INH TLM
0 1 4 8 1.5 95
Phospholipids
PE
CL
AC2PIM2/PI
AC1PIM2
Loading

Supplementary Figure 2: Analysis of apolar and polar phospholipids from GSK3011724A-treated *M. bovis* BCG. (**a**) *M. bovis* BCG cultures were treated with GSK3011724A, TLM and INH and labelled with [^{14}C]-acetate. Drug concentrations are displayed (μM). The total lipid extract was recovered and further solvent partitioned giving rise to apolar and polar phospholipids, which were analysed by autoradiography-TLC (10,000 cpm, equal counts). (**a**) 2D apolar lipid profiles using solvent systems A-C. (**b**) 1D polar phospholipid profile of polar phospholipids. PDIMs, phthiocerol dimycocerosates; TAG, triacylglycerol; MAT, multi-acylated trehaloses; F, fatty acids; GroM, monomycolylglycerol; PGL, phenolic glycolipid; PE, phosphatidylethanolamine; CL, cardiolipin; AC_2PIM_2, diacyl phosphatidylinositol dimannoside; PI, phosphatidylinositol; AC_1PIM_2, acyl phosphatidylinositol dimannoside.



Supplementary Figure 3: Inhibition of KasA-mediated condensation activity by GSK3011724A. KasA activity was monitored by the elongation of C_{16}-AcpM using [^{14}C]-malonyl-AcpM, which had been generated by the translocase reaction of FabD on holo-AcpM and [^{14}C]-malonyl-CoA. The condensation activity of KasA (0.25 μg) was assayed for 1.5 h in 0.1 % (v/v) DMSO with and without GSK3011724A (1 μM). In the absence of KasA, the enzyme was replaced by an equivalent volume of buffer. The results are based on the mean of triplicate reactions, with error bars representing the standard error.



Supplementary Figure 4: Surface Plasmon Resonance binding studies of KasA and GSK3011724A. Purified KasA was immobilised to a chip surface, and the binding of GSK3011724A was monitored. (**a**) Normalised signal from GSK3011724A titrations over immobilised KasA. The results are the mean of triplicate reactions, with error bars representing the standard error. (**b**) Representative sensorgrams of GSK3011724A titrations over immobilised KasA. Kinetic fits are shown.

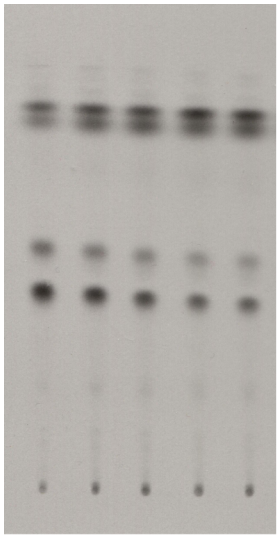


Figure 2a
(left panel)

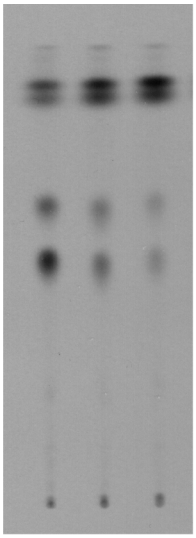


Figure 2a
(middle panel)

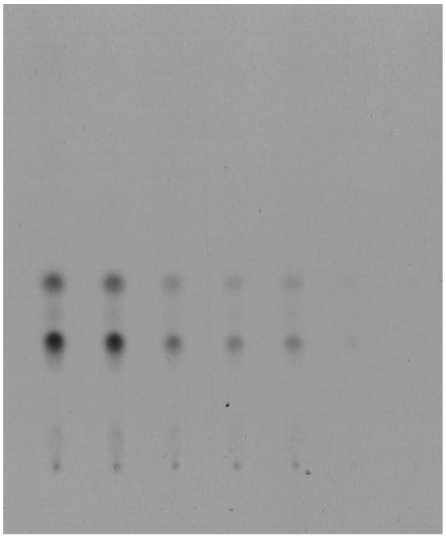
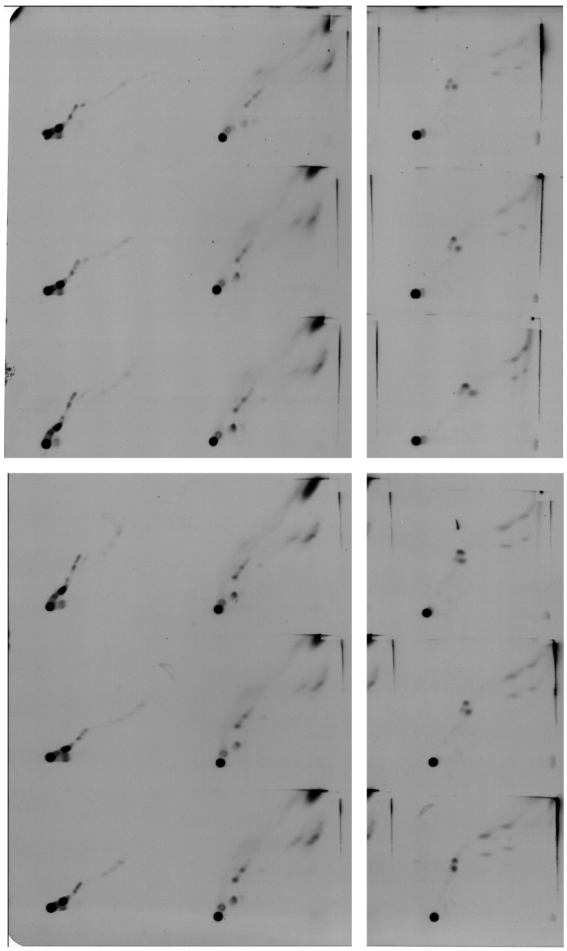


Figure 2a
(right panel)



Supplementary
Figure 2a

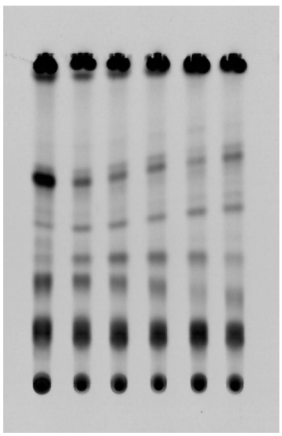
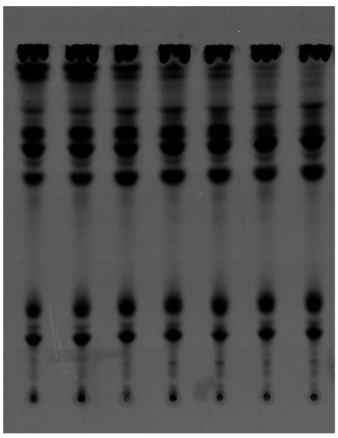
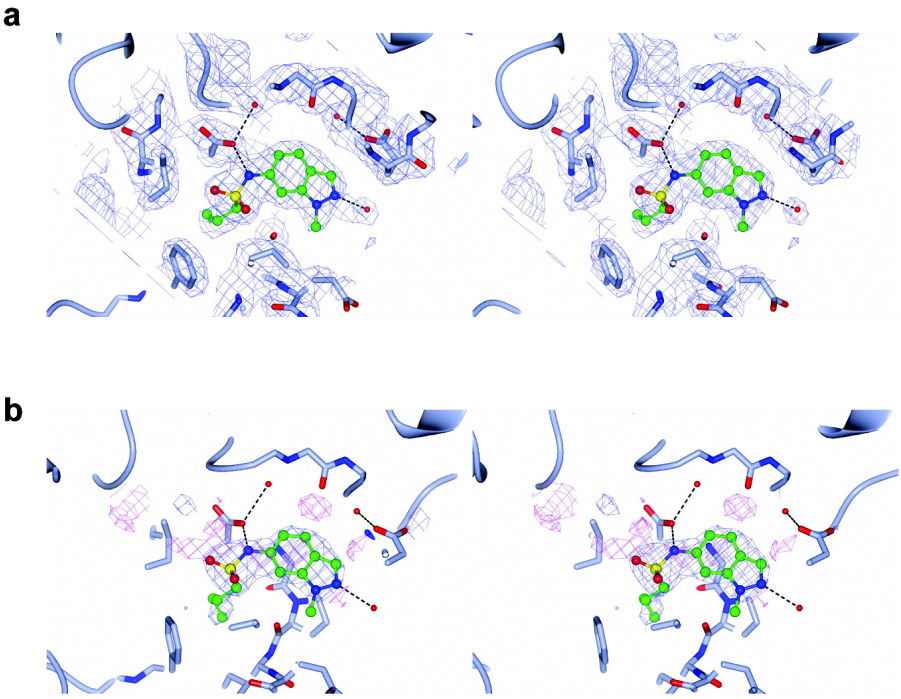


Figure 2b



Supplementary
Figure 2b

Supplementary Figure 5: Original scans for all TLC data.



Supplementary Figure 6: Stereo diagram of KasA complexed with GSK3011724A. (**a**) Fo-Fc density map for the ligand (contour level + 1.0 σ (blue)). (**b**) Fo-Fc density map for the ligand (contour level +/- 3.0 σ (blue/pink)).

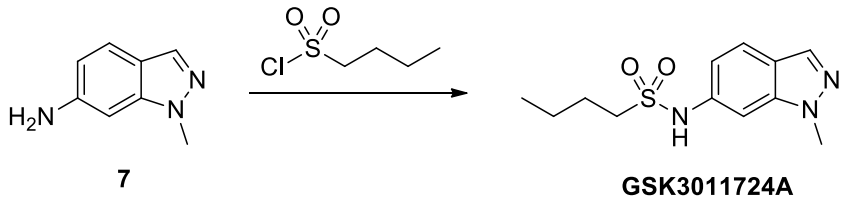
Supplementary Table

Assay	Mode	pXC_{50}
α1 nicotinic AChR	Opener	<4.3 (2)
α1 nicotinic AChR	Blocker	<4.3 (2)
Acetylcholinesterase	Inhibition	<4 (2)
Adenosine A_{2a}	Agonist	<4 (2)
Adrenergic α_{1B}	Antagonist	<4.6 (2)
Adrenergic α_{2C}	Agonist	<4 (2)
Adrenergic β_2	Agonist	<4 (2)
Adrenergic β_2	Antagonist	<4 (2)
AhR	Agonist	<4 (2)
Aurora B (STK12)	Antagonist	=5.1 (2)
Cannabinoid CB_2	Agonist	<4 (2)
COX-2	Blocker	<4 (2)
Dopamine D_1	Antagonist	<4 (2)
Dopamine D_2	Agonist	<4 (2)
Dopamine D_2	Antagonist	<4 (2)
hERG Ion Works	Antagonist	<4.2 (3)
Histamine H_1	Antagonist	<4.6 (2)
KCNQ1/minK	Blocker	<4.6 (2)
Kv1.5	Blocker	<4.3 (1)
LCK	Antagonist	<4.5 (2)
L-type Ca channel (CaV1.2)	Blocker	<4 (2)
Muscarine M_1	Agonist	<4.3 (2)
Muscarine M_1	Antagonist	<4.3 (2)
Muscarine M_2	Agonist	<4.3 (2)
Muscarine M_2	Antagonist	<4.3 (2)
NaV1.5	Blocker	<4 (2)
Neurokinin NK_1	Antagonist	<4.6 (2)
NMDA Channel (NR2B)	Blocker	<4.3 (2)
Norepinephrine	Antagonist	<4 (2)
OATP1B1	Inhibition	<4.3 (2)
Opioid μ	Agonist	<4 (2)
Opioid κ	Agonist	<4 (2)
PDE_{3A}	Inhibition	<4 (2)
PI3Kγ	Antagonist	<4.5 (1)
PXR	Agonist	<4.3 (2)
Serotonin $5HT_{1B}$	Agonist	<4 (2)
Serotonin $5HT_{1B}$	Antagonist	<4 (2)
Serotonin $5HT_{2A}$	Agonist	<4.6 (2)
Serotonin $5HT_{2A}$	Antagonist	<4.6 (2)
Serotonin $5HT_{2C}$	Agonist	<4.6 (2)
Serotonin $5HT_{2C}$	Antagonist	<4.6 (2)
Serotonin $5HT_3$	Opener	<4.3 (2)
Serotonin $5HT_3$	Blocker	<4.3 (2)
Vasopressin V_{1a}	Antagonist	<4.3 (2)

Supplementary Table 1: Off-target selectivity data for GSK3011724A.

Supplementary Methods

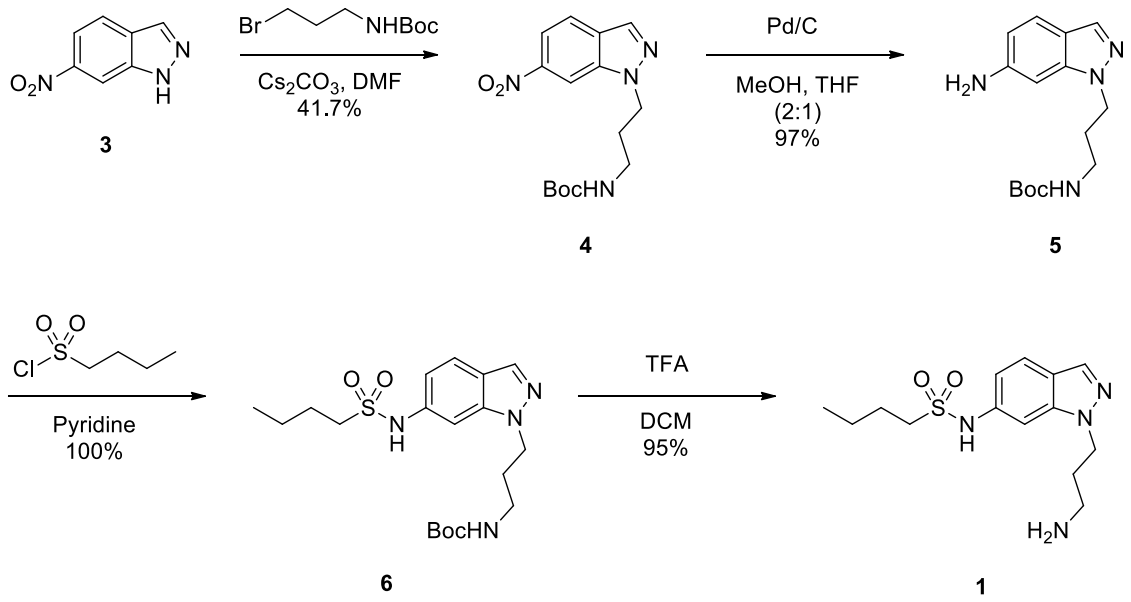
Synthesis and characterisation of compounds



Scheme 1: Synthesis of hit compound GSK3011724A

The synthesis of **GSK3011724A** was performed as follows. Analine **7** (929 mg, 6.31 mmol) was dissolved in dichloromethane (20 mL), and the solution was cooled to 0°C. Triethyl amine (2.64 mL, 18.94 mmol, 3 Eq.) was added followed by the sulfonyl chloride (1.95 mL, 15.78 mmol, 2.5 Eq.) over about 5 min, and then the mixture was allowed to warm to room temperate. After stirring for 3 h, complete consumption of the starting material was observed by LCMS. Methanol (20 mL) and 2 N aqueous sodium hydroxide (12.6 mL, 25.2 mmol, 4 Eq.) were then added to hydrolyse the bis-sulfonamide to the desired mono-sulfonamide, and stirred for 22 h at room temperature. The solution was concentrated to dryness, and the resulting mixture was diluted with dichloromethane and ammonium chloride. The aqueous layer was separated and extracted with dichloromethane, and the combined organic layers were dried over sodium sulfate. The resulting crude mixture was purified first by flash column chromatography (40 g cartridge) eluting with a 0-40 % iPrOH:cHex gradient. After concentration, the resulting coloured solid was re-crystallised from toluene to afford the desired product N-(1-methyl-1H-indazol-6-yl)butane-1-sulfonamide (**GSK3011724A**) as a white solid (1.004 g, 3.57 mmol, 57 %). mp:

120.6–122.2°C. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.95 (1H, s), 7.68 (1H, d, J = 8.6 Hz), 7.40 (1H, s), 7.15 (1H, s), 7.10 (1H, bs), 6.88 (1H, dt, J = 8.6, 1.5), 4.06 (3H, s), 3.13 (2H, m), 1.83 (2H, m), 1.42 (2H, m), 0.91 (3H, t, J = 7.3 Hz). ^{13}C NMR (100 MHz, DMSO-*d*6) δ ppm: 140.4, 137.3, 132.7, 122.2, 120.7, 115.0, 98.8, 50.6, 35.75, 25.6, 21.1, 13.9. LCMS (ES) [M]+H calculated for $C_{12}H_{17}N_3O_2S$ as 268.1120, found 268.1.



Scheme 2: Synthesis of linkable compound **1**

The synthesis of **1** was performed in a four-step reaction as follows: 6-nitro-1H-indazole (**3**; 203 mg, 1.24 mmol) was dissolved in dry N,N-Dimethylformamide (DMF) (4.1 mL). It was treated successively with caesium carbonate (1216 mg, 3.73 mmol, 3.0 Eq.) and tert-butyl (3-bromopropyl)carbamate (593 mg, 2.49 mmol, 2.0 Eq.). After 16 h the reaction was diluted with ethyl acetate, transferred to a separatory funnel, washed with brine solution 3 x, and the organic layer was dried with anhydrous sodium sulfate, filtered, and concentrated. The crude reaction mixture was purified using a 12 g silica column (30 mL.min^{-1}) on an Isco Brand auto column and was eluted with 40 % EtOAc/Hexanes to afford the desired product, tert-butyl (3-(6-

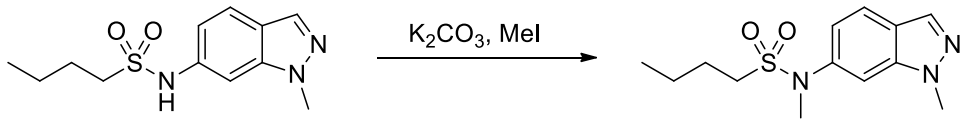
nitro-1H-indazol-1-yl)propyl)carbamate (**4**) in 41.7 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.41 (1H, s), 8.15 (1H, s), 8.05 (1H, d, J = 8.8 Hz), 7.86 (1H, d, J = 8.8 Hz), 4.78 (1H, br s), 4.57 (2H, t), 3.17 (2H, m), 2.19 (2H, m), 1.45 (9H, s); LCMS (ES) [M]+ calculated for $C_{15}H_{20}N_4O_4$ 320.1485, found 321.3 at 0.92 min.

Tert-butyl (3-(6-nitro-1H-indazol-1-yl)propyl)carbamate (58.9 mg, 0.184 mmol) was dissolved in Methanol (1.2 mL) and Tetrahydrofuran (THF) (0.61 mL) and then palladium on carbon (10 %) (11.74 mg, 0.110 mmol) was added. The atmosphere in the reaction flask was removed and replaced by hydrogen gas until the starting material had been consumed (2 h). The crude mixture was filtered through Celite, eluting with EtOAc, dried with anhydrous sodium sulfate and concentrated to afford the desired product, tert-butyl (3-(6-amino-1H-indazol-1-yl)propyl)carbamate (**5**) (51.6 mg, 0.178 mmol, 97 % yield), without further purification. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.81 (1H, s), 7.49 (1H, d, J = 8.3), 6.57 (1H, d, J = 8.5 Hz), 6.54 (1H, s), 4.85 (1H, br s), 4.32 (2H, t), 3.88 (2H, s br), 3.07 (2H, m), 2.05 (2H, s), 1.43 (9H, s); LCMS (ES) [M]+ calculated for $C_{15}H_{22}N_4O_2$ 290.1743, found 291.3 at 0.53 min.

Tert-butyl (3-(6-amino-1H-indazol-1-yl)propyl)carbamate (51.6 mg, 0.178 mmol) was dissolved in Pyridine (0.854 mL) and was cooled to 0ºC with an ice bath. Butane-1-sulfonyl chloride (35 µL, 0.270 mmol, 1.5 Eq.) was added drop wise. The mixture was stirred at 0°C and then allowed to reach room temperature. After 2 h, the reaction was quenched by addition of water and extracted with EtOAc, dried with $MgSO_4$ and solvent removed. The crude reaction mixture was diluted in EtOAc, washed subsequently with 1.0 N HCl and then brine and the organic solvent was removed *in vacuo*. The crude reaction mixture was dissolved on silica and tert-butyl (3-(6-

(butylsulfonamido)-1H-indazol-1-yl)propyl)carbamate (**6**) was purified using a 4 g silica column (18 mL.min^{-1}) on an Isco Brand auto column, eluted with (30 %) EtOAc/Hexanes. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.96 (1H, s), 7.68 (1H, d, J = 8.5), 7.42 (1H, s), 7.19 (1H, s), 6.97 (1H, br s), 4.88 (1H, s br), 4.44 (2H, t), 3.12 (4H, m), 2.05 (2H, m), 1.84 (2H, m), 1.44 (11H, m), 0.91 (3H, t); LCMS (ES) [M]+ calculated for $C_{19}H_{30}N_4O_4S$ 410.1988, found 411.4 at 0.83 min.

Tert-butyl (3-(6-(butylsulfonamido)-1H-indazol-1-yl)propyl)carbamate (47.6 mg, 0.116 mmol) was dissolved in DCM (1.739 mL), and Trifluoroacetic acid (TFA) (0.580 mL) was added and stirred for 1 h at room temperature. The reaction was condensed to dryness and the resulting oil was dissolved in MeOH and loaded onto an Isolute SCX column. The material was washed with MeOH and then the desired compound was eluted off of the column using 7.0 N Ammonia in MeOH. The solvent was removed to afford the desired product, N-(1-(3-aminopropyl)-1H-indazol-6-yl)butane-1-sulfonamide (**1**) (34.3 mg, 0.110 mmol, 95 % yield), as an oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.96 (1H, s), 7.67 (1H, d, J = 8.5), 7.45 (1H, s), 6.92 (1H, d, J=8.5), 4.48 (2H, m), 3.26 (3H, s br), 3.13 (2H, m), 2.72 (2H, m), 2.06 (2H, m), 1.83 (2H, m), 1.42 (2H, m), 0.90 (3H, t); LCMS (ES) [M]+ calculated for $C_{14}H_{22}N_4O_2S$ 310.1463, found 311.3 at 0.42 min.



GSK3011724A **2**

Scheme 3: Synthesis of inactive compound **2**

The synthesis of **2** was performed as follows. **GSK3011724A** (110 mg, 0.411 mmol) was dissolved in N,N-dimethylformamide (10 mL), and the solution was cooled to 0°C. Potassium carbonate (171 mg, 1.234 mmol, 3 Eq.) was added followed by iodomethane (100 µL, 1.606 mmol, 4.0 Eq.), and then the mixture was allowed to warm to room temperate and left stirring overnight. The reaction was diluted with water and ethyl acetate. The organic phase was washed several times with aqueous ammonium chloride, dried over sodium sulfate, and concentrated under reduced pressure. The resulting crude mixture was purified first by flash column chromatography (4 g cartridge) eluting with a 0-40 % EtOAc:cHex gradient and concentrated under reduced pressure to afford the desired product N-methyl-N-(1-methyl-1H-indazol-6-yl)butane-1-sulfonamide (**2**) as a white solid (92.5 mg, 0.329 mmol, 80 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.98 (1H, s), 7.73 (1H, d, J = 8.6 Hz), 7.49 (1H, s), 7.15 (1H, s), 7.14 (1H, dd, J = 8.6, 1.8), 4.07 (3H, s), 3.43 (3H, s), 3.02 (2H, m), 1.83 (2H, m), 1.43 (2H, m), 0.93 (3H, t, J = 7.3 Hz). LCMS (ES) [M]+H calculated for $C_{13}H_{19}N_3O_2S$ as 282.1276, found 282.1.

Surface Plasmon Resonance studies

Surface Plasmon Resonance (SPR) binding studies were performed using a Biacore T200 instrument (GE Healthcare) and a NiHC1000m chip (XanTec bioanalytics GmbH, Dusseldorf). Chip preparation and binding analyses were performed at 25°C, with 10 mM HEPES pH 7.4, 150 mM NaCl, 200 µM TCEP, 0.005 % Tween 20, 1 % DMSO as running buffer. After chip conditioning (120 s injection of 0.5 M EDTA, 60 s injection of 500 µM $NiCl_2$, both at 10 $\mu l.min^{-1}$), KasA was captured by flowing protein (50 $\mu g.ml^{-1}$ in running buffer) over sample flow-cells for 5-7 min, capturing

10,000 – 15,000 resonance units (RU) of protein. Flow-cell 1 remained unmodified and was used as the reference channel. Binding analyses for GSK3011724A were performed at 30 µl.min^{-1}, using 120 s sample injection times with 1,800 s dissociation. GSK3011724A was diluted to 1 µM in running buffer, then serially diluted 7 times in 3-fold steps, and samples were injected starting with the most dilute. Analyses were repeated on the same surface. Initial dose-responses were discarded due to baseline decay. Only data from matured KasA surfaces, where there was a suitable baseline, were analysed. For each dose-response data set, responses at the end of the association phase were normalised to 0-100%, then combined with replicates and the data fitted with a 1:1 binding model.

***In vitro* death curve assay**

M. tuberculosis H37Rv was cultured at 37^{o}C to OD_{600nm} 0.1-0.8 in Middlebrook 7H9 broth supplemented with Middlebrook ADC and 0.025 % (v/v) tyloxapol. The culture was diluted to 5 x 10^{5} CFU.ml^{-1} (10 mL). Cells were treated with 10 x MIC of the following drugs: GSK3011724A (5 µM), moxifloxacin (1.5 µM), linezolid (17.8 µM). A no drug control for positive growth was also included. Cultures were grown in triplicate. A 100 µL aliquot was sampled at 0, 2, 4, 7 and 14 days and the cells were plated onto Middlebrook 7H10 solid media supplemented with Middlebrook OADC. Dilutions corresponding to 10, 100 and 1000 CFU were performed in PBS containing 0.025 % (v/v) tyloxapol, and were plated in triplicate.